THE CORRESPONDENCE THEOREM

A THESIS

Presented to

The Faculty of the Graduate Division

by

Frank Harold Morgan

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Applied Mathematics

Georgia Institute of Technology

October, 1964

THE CORRESPONDENCE THEOREM

Approved:

Date approved by Chairman: 8 October 1964

ACKNOWLEDGMENTS

I wish to express my sincere appreciation to Dr. James W. Walker, my thesis advisor, for his guidance and help in the preparation of this thesis. I also wish to express my gratitude to Dr. Eric R. Immel and Dr. David L. Finn for their reading of the manuscript and their helpful suggestions in improving it. I also wish to thank Mrs. Peggy Weldon for her excellent typing of the thesis.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

The correspondence between probability measures - more generally Lebesgue-Stieltjes measures - and distribution functions is essential in studies of mathematical statistics and probability theory. This correspondence, although rarely mentioned in elementary works, is usually given at more advanced levels (cf. CRAMÉR [1], pp. 80-81; KOLMOGOROV, pp. 25-26; LOÈVE, pp. 96-98; WILKS [1], pp. 9-10; and WILKS [2], pp. 31-33, 39-41). Loève outlines the proof for the one-dimensional case, and states, perhaps misleadingly, that the generalization to the multidimensional case is immediate. CRAMÉR [1] and KOLMOGOROV state the correspondence and refer to LEBESGUE (pp. 152-156, 168) and RADON for the proofs. WILKS [1] states, but does not prove, weaker versions of the one- and two-dimensional cases.

Lebesgue seems to have discovered the one-dimensional case, or rather a result similar to it, prior to 1903. Radon gives somewhat of a generalization to closed intervals about the origin in E^2 (two-dimensional euclidean space) but does not prove this generalization for any higher dimensional case.

More recently the correspondence has been investigated from a measure-theoretic standpoint without any direct appeal to the notion of distribution functions. For example KAMKE (pp. 41-47) arrives at the notion of some types of measures via the approach of finitely additive functions defined on intervals, and VON NEUMANN (pp. 160-170) proves a powerful theorem relating total additivity to the monotone property possessed by the difference operator but never explicitly gives the correspondence itself.

Probabilists and statisticians tend to present cursory developments of the correspondence theorem while measure-theoretists, although tending to neglect the correspondence itself, nevertheless obtain pertinent results in more abstract settings. Two purposes of this paper are to relate these probabilistic and statistical discussions to the similar and more general measure-theoretic results and to provide an easily readable transition from the former to the latter.

The difference operator $\Delta_{k_1 \dots k_n}$ for functions on E^N (N-dimensional euclidean space) and left-continuity are discussed in Chapter II. Theorems concerning the permutation of the differencing subscripts in the operator and the representation of the operator as a finite sum are stated and proved. A necessary and sufficient condition for left-continuity utilizing the difference operator is given.

In Chapter III the correspondence theorem in one dimension is stated in three parts and proved in such a way that the transition to the N-dimensional case can be made as easily as possible.

The N-dimensional case and the difficulties in generalizing the one-dimensional case are discussed in Chapter IV.

In Chapter V the superiority of some of von Neumann's terminology over that of Loève and some of von Neumann's measure-theoretic generalizations are discussed. It is shown that, although von Neumann does not state the correspondence theorem, portions of it follow from some of his theorems.

The following notations are common throughout the text: E_j and E^j denote, respectively, one-dimensional and j-dimensional euclidean space; the union of disjoint sets A_j is denoted by ΣA_j; the empty set is denoted by $\emptyset$; and the symbol ∎ is the "Halmos Finality Symbol" which indicates the end of a proof.

Certain well-known theorems are cited and are given here:

<u>Carathéodory Extension Theorem</u>: A measure μ on a field $\mathcal{C}$ can be extended to a measure on the minimal σ-field over $\mathcal{C}$. If in addition μ is σ-finite, then the extension measure is unique and σ-finite.

<u>Heine-Borel Theorem</u>: Let F be an open cover of a closed and bounded set A in E^N. Then a finite subcollection of F also covers A.

<u>Continuity Theorem for Measures</u>: If $\{A_n\}$ is a sequence of sets such that $A_i \supset A_{i+1}$, $i = 1, 2, \ldots,$ and $\bigcap_{n=1}^{\infty} A_n = \emptyset$, and if μ is a measure on a field containing the A_n, then $\lim_{n \to \infty} \mu A_n = 0$, provided $\mu A_r < \infty$ for some r.

CHAPTER II

DIFFERENCING AND CONTINUITY FROM THE LEFT

In this chapter a difference operator for functions on E^N (N-dimensional euclidean space) is defined. Theorems concerning its properties, its representation as a finite sum, and its relation to continuity are proved. The usual notion of continuity from the left in E^1 is generalized to E^N and a theorem relating continuity from the left to the difference operator is partially proved. The remainder of the proof is given in Chapter IV.

Unless it is otherwise stated, g will be a real-valued function defined on an open set $S \subset E^N$, and $n \leq N$.

Differencing

<u>Definition 2.1</u>: The difference operator $\Delta_{k_1 \ldots k_n}$ is defined recursively as follows:

$$\Delta_{k_1} g(x; \lambda) = h_\lambda(x) = g(x + \lambda u_{k_1}) - g(x)$$

where u_i is a unit vector with unity in the i^{th} position; and, for $n \geq 2$,

$$\Delta_{k_1 \ldots k_n} g(x; \lambda_1, \ldots, \lambda_n) = \Delta_{k_n} h_{\lambda_1 \ldots \lambda_{n-1}}(x; \lambda_n) = h_{\lambda_1 \ldots \lambda_n}(x).$$

Each of the integers $k_1, k_2, \ldots, k_n$ is to be in $\{1, 2, \ldots, N\}$.

For convenience, $\Delta_{k_1 \ldots k_n} g(x; \lambda_1, \ldots, \lambda_n)$ will be written $\Delta_{k_1 \ldots k_n} g(x; \lambda)$ and when g is a function of one variable, $\Delta_1 g(x; \lambda)$ will be written $\Delta g(x; \lambda)$.

<u>Theorem 2.2</u>: $\Delta_{k_1\ldots k_n} g(x; \lambda_1,\ldots,\lambda_n) = \Delta_{1\ldots n} g(x; \lambda'_1,\ldots,\lambda'_n)$ where $(k_1,\ldots,k_n)$ is any permutation of $(1,\ldots,n)$ and $(\lambda_1,\ldots,\lambda_n)$ is the corresponding permutation of $(\lambda'_1,\ldots,\lambda'_n)$.

<u>Proof</u>. The assertion is obviously true for $n = 1$. Assume it true for all $r \leq n - 1$. There are two cases.

Case 1. Suppose $k_n = n$. By definition

$$\Delta_{1\ldots n} g(x; \lambda') = \Delta_n h_{\lambda'_1\ldots\lambda'_{n-1}} (x; \lambda'_n).$$

By the inductive assumption

$$h_{\lambda'_1\ldots\lambda'_{n-1}}(x) = \Delta_{1\ldots n-1} g(x; \lambda'_1,\ldots,\lambda'_{n-1}) =$$

$$= \Delta_{k_1\ldots k_{n-1}} g(x;\lambda_1,\ldots,\lambda_{n-1}) = h_{\lambda_1\ldots\lambda_{n-1}}(x).$$

Hence, according to the definition,

$$\Delta_{1\ldots n} g(x; \lambda') = \Delta_n h_{\lambda'_1\ldots\lambda'_{n-1}} (x; \lambda'_n) =$$

$$= h_{\lambda'_1\ldots\lambda'_{n-1}} (x + \lambda'_n u_n) - h_{\lambda'_1\ldots\lambda'_{n-1}}(x) =$$

$$= h_{\lambda_1\ldots\lambda_{n-1}} (x + \lambda'_n u_n) - h_{\lambda_1\ldots\lambda_{n-1}}(x) =$$

$$= \Delta_n h_{\lambda_1\ldots\lambda_{n-1}}(x; \lambda'_n) = \Delta_{k_1\ldots k_{n-1}n} g(x; \lambda_1,\ldots,\lambda_{n-1},\lambda'_n)$$

The last expression is $\Delta_{k_1\ldots k_n} g(x; \lambda)$ since it is assumed that $k_n = n$.

Case 2. Suppose that $k_r = n$, where $r \neq n$. Then

$$\Delta_{k_1\ldots k_n} g(x; \lambda) = \Delta_{k_n} h_{\lambda_1\ldots\lambda_{n-1}} g(x; \lambda_n) =$$

$$= h_{\lambda_1\dots\lambda_{n-1}}(x + \lambda_n u_{k_n}) - h_{\lambda_1\dots\lambda_{n-1}}(x) =$$

$$= \Delta_{k_1\dots k_{n-1}} g(x + \lambda_n u_{k_n}; \lambda_1,\dots,\lambda_{n-1}) - \Delta_{k_1\dots k_{n-1}} g(x; \lambda_1,\dots,\lambda_{n-1}) =$$

$$= \Delta_{k_{n-1}} h_{\lambda_1\dots\lambda_{n-2}}(x + \lambda_n u_{k_n}; \lambda_{n-1}) - \Delta_{k_{n-1}} h_{\lambda_1\dots\lambda_{n-2}}(x; \lambda_{n-1}) =$$

$$= h_{\lambda_1\dots\lambda_{n-2}}(x + \lambda_n u_{k_n} + \lambda_{n-1} u_{k_{n-1}}) - h_{\lambda_1\dots\lambda_{n-2}}(x+\lambda_n u_{k_n}) +$$

$$- h_{\lambda_1\dots\lambda_{n-2}}(x + \lambda_{n-1} u_{k_{n-1}}) + h_{\lambda_1\dots\lambda_{n-2}}(x) =$$

$$= \Delta_{k_n} h_{\lambda_1\dots\lambda_{n-2}}(x + \lambda_{n-1} u_{k_{n-1}}; \lambda_n) - \Delta_{k_n} h_{\lambda_1\dots\lambda_{n-2}}(x; \lambda_n) =$$

$$= \Delta_{k_1\dots k_{n-2} k_n} g(x+\lambda_{n-1} u_{k_{n-1}}; \lambda_1,\dots,\lambda_{n-2}, \lambda_n) +$$

$$- \Delta_{k_1\dots k_{n-2} k_n} g(x; \lambda_1,\dots,\lambda_{n-2}, \lambda_n) =$$

$$= h_{\lambda_1\dots\lambda_{n-2}\lambda_n}(x + \lambda_{n-1} u_{k_{n-1}}) - h_{\lambda_1\dots\lambda_{n-2}\lambda_n}(x) =$$

$$= \Delta_{k_{n-1}} h_{\lambda_1\dots\lambda_{n-2}\lambda_n}(x; \lambda_{n-1}) = \Delta_{k_1\dots k_{n-2} k_n k_{n-1}} g(x; \lambda_1,\dots,\lambda_{n-2}, \lambda_n, \lambda_{n-1}).$$

Also

$$\Delta_{k_1\dots k_n} g(x; \lambda) = \Delta_{k_n} h_{\lambda_1\dots\lambda_{n-1}}(x; \lambda_n) =$$

$$= h_{\lambda_1\dots\lambda_{n-1}}(x + \lambda_n u_{k_n}) - h_{\lambda_1\dots\lambda_{n-1}}(x) =$$

$$= \Delta_{k_1\dots k_{n-1}} g(x + \lambda_n u_{k_n}; \lambda_1,\dots,\lambda_{n-1}) +$$

$$- \Delta_{k_1\dots k_{n-1}} g(x; \lambda_1,\dots, \lambda_{n-1}).$$

By the inductive assumption this last difference is

$$\Delta_{k'_1\ldots k'_{n-1}}\, g(x + \lambda_n\ u_{k_n};\ \lambda'_1,\ldots,\lambda'_{n-1}) +$$

$$-\ \Delta_{k'_1\ldots k'_{n-1}}\, g(x;\ \lambda'_1,\ldots,\lambda'_{n-1}),$$

where $k'_{n-1} = n$, where $(k'_1,\ldots,k'_{n-2})$ is the same permutation of $(k_1,\ldots,k_{n-2})$ as $(\lambda'_1,\ldots,\lambda'_{n-2})$ is of $(\lambda_1,\ldots,\lambda_{n-2})$, and where $\lambda'_{n-1} = \lambda_{k_r}$. Thus

$$\Delta_{k_1\ldots k_n}\, g(x;\ \lambda) = \Delta_{k'_1\ldots k'_{n-1}k_n}\, g(x;\ \lambda'),$$

where $k'_{n-1} = n$, and where the subscripts and increments of the left member are appropriately permuted.

So far it has been shown that

$$\Delta_{k_1\ldots k_n}\ g(x;\ \lambda) = \Delta_{k'_1\ldots k'_{n-2}\,n\,k_n}\ g(x;\ \lambda'_1,\ldots,\lambda'_{n-1},\lambda_n) =$$

$$= \Delta_{k'_1\ldots k'_{n-2}k_n n}\ g(x;\ \lambda'_1,\ldots,\lambda'_{n-2},\lambda_n,\lambda'_{n-1}).$$

The assertion follows from the result of Case 1. This completes the proof. ∎

Corollary 2.3: $\Delta_{k'_1\ldots k'_n}\ g(x;\ \lambda'_1,\ldots,\lambda'_n) = \Delta_{k_1\ldots k_n}\ g(x;\ \lambda_1,\ldots,\lambda_n)$, where $(k'_1,\ k'_2,\ldots,k'_n)$ is a permutation of $(k_1,k_2,\ldots,k_n)$, and $(\lambda'_1,\lambda'_2,\ldots,\lambda'_n)$ is the corresponding permutation of $(\lambda_1,\lambda_2,\ldots,\lambda_n)$.

Proof. Relabel the variables $x_1,\ldots,x_N$ so that x'_i is x_{k_i} for $i = 1,2,\ldots,n$. Define G by $G(x'_1,\ldots,x'_N) = g(x_1,\ldots,x_N)$. Then

$\Delta_{k_1 \ldots k_n} g(x; \lambda) = \Delta_{1 \ldots n} G(x'; \lambda)$ for all relevant x and all λ.
But by Theorem 2.2, $\Delta_{1 \ldots n} G(x'; \lambda) = \Delta_{r_1 \ldots r_n} G(x'; \lambda')$, where $(r_1, r_2, \ldots, r_n)$ is the same permutation of $(1,2,\ldots,n)$ as $(k_1', k_2', \ldots, k_n')$ is of $(k_1, k_2, \ldots, k_n)$, and $\Delta_{r_1 \ldots r_n} G(x'; \lambda') = \Delta_{k_1' \ldots k_n'} g(x; \lambda')$.
The conclusion follows from the three equalities. ∎

Theorem 2.4:

$$\Delta_{1 \ldots n} g(x; \lambda) = \sum_{j=o}^{n} (-1)^{n-j} \sum_{\binom{n}{j}} g(x + \sum_{s=1}^{j} \lambda_{i_s} u_{i_s})$$

with the understanding that $\sum_{\binom{n}{j}}$ denotes summation over all subsets $\{i_1, i_2, \ldots, i_j\}$ of $\{1,2,\ldots,n\}$, and that for $j = 0$ its value is $g(x)$.

Proof: The assertion is obvious for $n = 1$. Assume it true for $n = k-1$.
By Definition 2.1

$$\Delta_{1 \ldots k} g(x; \lambda_1, \ldots, \lambda_k) = \Delta_{1 \ldots k-1} g(x + \lambda_k u_k; \lambda_1, \ldots, \lambda_{k-1}) + \\ - \Delta_{1 \ldots k-1} g(x; \lambda_1, \ldots, \lambda_{k-1}).$$

By the inductive assumption the last two differences can be written as sums so that

$$\Delta_{1 \ldots k} g(x; \lambda_1, \ldots, \lambda_k) = \sum_{j=o}^{k-1} (-1)^{k-1-j} \sum_{\binom{k-1}{j}} g(x + \lambda_k u_k + \sum_{s=1}^{j} \lambda_{i_s} u_{i_s}) + \\ - \sum_{j=o}^{k-1} (-1)^{k-1-j} \sum_{\binom{k-1}{j}} g(x + \sum_{s=1}^{j} \lambda_{i_s} u_{i_s}) .$$

The difference of sums will be transformed into the desired result.

$$\Delta_{1\ldots k}\, g(x;\lambda_1,\ldots,\lambda_k) =$$

$$= \sum_{j=o}^{k-2} (-1)^{k-1-j} \sum_{\binom{k-1}{j}} g(x+\lambda_k u_k + \sum_{s=1}^{j} \lambda_{i_s} u_{i_s}) +$$

$$- \sum_{j=1}^{k-1} (-1)^{k-1-j} \sum_{\binom{k-1}{j}} g(x + \sum_{s=1}^{j} \lambda_{i_s} u_{i_s}) +$$

$$+ g(x+\sum_{s=1}^{k} \lambda_s u_s) + (-1)^k g(x) =$$

$$= \sum_{j=1}^{k-1} (-1)^{k-j} \left\{ \sum_{\binom{k-1}{j-1}} g(x+\lambda_k u_k + \sum_{s=1}^{j-1} \lambda_{i_s} u_{i_s}) + \sum_{\binom{k-1}{j}} g(x+\sum_{s=1}^{j} \lambda_{i_s} u_{i_s}) \right\} +$$

$$+ g(x+\sum_{s=1}^{k} \lambda_s u_s) + (-1)^k g(x) .$$

Since the sum over $\binom{k}{j}$ is the sum of all values $g(c_1,\ldots,c_k, x_{k+1},\ldots,x_N)$ for which exactly j of the c's are of the form $x_i+\lambda_i$, with the remaining ones of the form x_i, it can be written as the sum of two quantities: the first being the sum of those values for which $c_k = x_k + \lambda_k$ and the second being the sum of those values for which $c_k = x_k$. These two sums are the ones in braces. Their sum is therefore

$$\sum_{\binom{k}{j}} g(x + \sum_{s=1}^{j} \lambda_{i_s} u_{i_s}).$$

Hence

$$\Delta_{1\ldots k}\, g(x;\lambda) = \sum_{j=1}^{k-1} (-1)^{k-j} \sum_{\binom{k}{j}} g(x + \sum_{s=1}^{j} \lambda_{i_s} u_{i_s}) + g(x + \sum_{s=1}^{k} \lambda_{i_s} u_{i_s}) +$$

$$+ (-1)^k g(x) = \sum_{j=o}^{k} (-1)^{k-j} \sum_{\binom{k}{j}} g(x + \sum_{s=1}^{j} \lambda_{i_s} u_{i_s}).$$

This completes the proof. ∎

<u>Corollary 2.5</u>: $\Delta_{k_1 \ldots k_n} g(x;\lambda_1,\ldots,\lambda_n) = \sum_{j=o}^{n} (-1)^{n-j} \sum_{\binom{n}{j}} g(x + \sum_{s=1}^{j} \lambda_{i_s} u_{k_{i_s}})$.

<u>Proof</u>: Relabel the x_i as in the proof of Corollary 2.3. ∎

<u>Theorem 2.6</u>: A necessary condition that g be continuous at x is that $\lim_{\substack{\lambda \to o \\ \lambda \in E^r}} \Delta_{1\ldots r}\, g(x;\lambda) = 0$ for $r = 1,2,\ldots,n$.

<u>Proof</u>: $\lim_{\lambda \to o} \Delta_{1\ldots r}\, g(x;\lambda) = \sum_{j=o}^{r} (-1)^{r-j} \sum_{\binom{r}{j}} \lim_{\lambda \to o} g(x + \sum_{s=1}^{j} \lambda_{i_s} u_{i_s}) =$

$$= g(x) \sum_{j=o}^{r} (-1)^{r-j} \binom{r}{j} = g(x)(1-1)^r = 0. \;\blacksquare$$

<u>Theorem 2.7</u>: If $\lim_{\substack{\lambda \to o \\ \lambda \in E^r}} \Delta_{k_1 \ldots k_r}\, g(x;\lambda) = 0$ for $r = 1,2,\ldots,n$ for every $k_1 < k_2 < \ldots < k_r$ where $k_j \in \{1,2,\ldots,n\}$, then $\lim_{\substack{\lambda \to o \\ \lambda \in E^n}} g(x + \sum_{s=1}^{n} \lambda_s u_s) = g(x)$.

Proof: For $n = 1$ there is nothing to prove since both limit assertions are the same. Assume the theorem true for $n \leq q$. Then

$$\Delta_{k_1 \ldots k_{q+1}} g(x; \lambda) =$$

$$= \Delta_{k_1 \ldots k_q} g(x + \lambda_{q+1} u_{q+1}; \lambda_1, \ldots, \lambda_q) - \Delta_{k_1 \ldots k_q} g(x; \lambda_1, \ldots, \lambda_q) =$$

$$= \sum_{j=o}^{q} (-1)^{q-j} \sum_{\binom{q}{j}} g(x + \lambda_{q+1} u_{q+1} + \sum_{s=1}^{j} \lambda_{i_s} u_{k_{i_s}}) - \sum_{j=o}^{q} (-1)^{q-j} \sum_{\binom{q}{j}} g(x + \sum_{s=1}^{j} \lambda_{i_s} u_{k_{i_s}}) =$$

$$= g(x + \sum_{s=1}^{q+1} \lambda_s u_s) + \sum_{j=o}^{q-1} (-1)^{q-j} \sum_{\binom{q}{j}} g(x + \lambda_{q+1} u_{q+1} + \sum_{s=1}^{j} \lambda_{i_s} u_{k_{i_s}}) +$$

$$- \sum_{j=o}^{q} (-1)^{q-j} \sum_{\binom{q}{j}} g(x + \sum_{s=1}^{j} \lambda_{i_s} u_{k_{i_s}}) .$$

The function values in each of the last two sums over $\binom{q}{j}$ are of the form $g(x + \sum_{s=1}^{r} \lambda_{k_s} u_{k_s})$ where $r \leq q$ so that by the inductive assumption each of these tends to $g(x)$. Hence

$$0 = \lim_{\lambda \to o} \Delta_{k_1 \ldots k_{q+1}} g(x; \lambda) = \lim_{\lambda \to o} g(x + \sum_{s=1}^{q+1} \lambda_s u_s) +$$

$$+ g(x) \left\{ \sum_{j=o}^{q-1} (-1)^{q-j} \binom{q}{j} - \sum_{j=o}^{q} (-1)^{q-j} \binom{q}{j} \right\} =$$

$$= \lim_{\lambda \to o} g(x + \sum_{s=1}^{\delta+1} \lambda_s u_s) - g(x) .$$

The conclusion follows.

Theorem 2.7 gives a sufficient condition for continuity of g. Furthermore, a slight modification of the proof of Theorem 2.6 shows that this condition is also necessary. The difference between this condition and the one of Theorem 2.6 is illustrated by the following

<u>Theorem 2.8</u>: The condition in Theorem 2.6 is not sufficient.

<u>Proof</u>: Consider the function $g(x_1, x_2)$ which is zero if $x_2 = 0$ and unity otherwise. Then

$$\Delta_1 g(0;\ \lambda_1) \equiv 0 \text{ and } \Delta_{12}\ g(0;\ \lambda_1,\ \lambda_2) \equiv 0\ .$$

but g is discontinuous at the origin.

Continuity from the Left

<u>Definition 2.9</u>: For any $x \epsilon E^N$ and any $(t_1, t_2, \ldots, t_k)$, where $t_1 < t_2 < \ldots < t_k$ and $t_j \in 1, 2, \ldots, N$ for $j = 1, 2, \ldots, k$, write $\overline{x}_{t_k} = (x_{t_1}, x_{t_2}, \ldots, x_{t_k})$. g is said to be <u>continuous from the left</u> at x provided that, for every $\{t_1, t_2, \ldots, t_k\} \subset \{1, 2, \ldots, N\}$, with $t_1 < t_2 < \ldots < t_k$,

$$\lim_{\overline{a}_{t_k} \to \overline{x}_{t_k}} \Delta_{1 \ldots N}\, g(a;\ x-a) = 0$$

where, for each relevant t_j, $-\infty < a_{t_j} \leq x_{t_j} < \infty$.

It is clear that for N = 1 the above definition describes the usual notion of continuity from the left. The following theorem considerably simplifies the notion of continuity from the left if the function g has a monotone property ($\Delta_{1 \ldots N}\, g \geq 0$) .

Theorem 2.10: Suppose g is defined throughout E^N and that $\Delta_{1 \ldots N}\, g(a;\ x - a) \geq 0$ for each $a = (a_1, a_2, \ldots, a_N)$ such that $a_i \leq x_i$, $i = 1,2,\ldots,N$. A necessary and sufficient condition that g be continuous from the left at x is that

$$\lim_{a_i \to x_i} \Delta_{1 \ldots N}\, g(a;\ x - a) = 0$$

for $i = 1,2,\ldots,N$.

A Discussion of the Proof. The condition is obviously necessary; its sufficiency follows from the finite additivity of $\Delta_{1 \ldots N}$ on the class of all intervals half-open on the right. The proof is therefore delayed until this property is established.

CHAPTER III

THE CORRESPONDENCE THEOREM IN ONE DIMENSION

In this chapter the concepts of "distribution function" and "Lebesgue-Stieltjes measure" will be defined in E^1, and the correspondence theorem in one dimension will be stated and proved. Assumed is an acquaintance with the elementary concepts of the theory of functions of a real variable, especially the material in Part One of LOÈVE.

Definition 3.1: A real-valued function F on E^1 is said to be a distribution function in case it is finite, nondecreasing, and continuous from the left throughout E^1.

Definition 3.2: A measure defined on a class S of sets in a euclidean space is said to be a Lebesgue-Stieltjes measure on S if it is finite on any finite interval in S.

The Statement of the Theorem

The theorem is stated in three parts.

Part I: Let $\mathcal{B}_I = \left\{[a,b) \mid -\infty < a \leq b < \infty\right\}$. A Lebesgue-Stieltjes measure μ on $\mathcal{B}_I$ extends uniquely to a Lebesgue-Stieltjes measure on the class $\mathcal{B}_0 = \left\{\sum_{k=1}^{n} I_k + I^c \mid I_k \in \mathcal{B}_I;\ I \in \mathcal{B}_I \text{ or } I = E^1\right\}$, a field over $\mathcal{B}_I$; the Borel field $\mathcal{B}$ is the minimal σ-field over $\mathcal{B}_0$; and μ on $\mathcal{B}_0$ extends uniquely to a Lebesgue-Stieltjes measure on

$\mathcal{B}_\mu = \left\{A \mid A = B + N, \text{ where } B \varepsilon \mathcal{B} \text{ and } N \subset Q \varepsilon \mathcal{B} \text{ with } \mu Q = 0\right\}$, the completion of the Borel field for μ.

Part II: The relation, ~, on the class $\mathcal{F}$ of distribution functions defined by

$$F \sim G \text{ if, and only if, } \Delta F(x; \lambda) = \Delta G(x; \lambda)$$

for all $x \varepsilon E^1$ and all $\lambda \varepsilon E^1$ is an equivalence relation on $\mathcal{F}$. The relation

$$\mu[a, b) = \Delta F(a; b-a) ,$$

where $-\infty < a \leq b < \infty$, defines a finite measure μ on $\mathcal{B}_I$.

Part III: The function φ which assigns to each equivalence class $\mathcal{F}_F$ (represented by F) the measure of Part II establishes a one-to-one correspondence between the equivalence classes and the class of Lebesgue-Stieltjes measures on $\mathcal{B}_I$.

The Proof of Part I

It will first be shown that the class $\mathcal{B}_o$ of Part I is a field and that $\mathcal{B}$ is minimal over $\mathcal{B}_o$. The extensions from $\mathcal{B}_I$ to $\mathcal{B}_o$ and from $\mathcal{B}$ to $\mathcal{B}_\mu$ are accomplished in a natural way and the extension from $\mathcal{B}_o$ to $\mathcal{B}$ is guaranteed by the Carathéodory extension theorem.

Theorem 3.3: The class $\mathcal{B}_o$ is a field.

Proof: Let $A = \sum_{k=1}^{n} I_k + I^c$ and $B = \sum_{k=1}^{m} J_k + J^c$ be arbitrary sets in $\mathcal{B}_o$. By definition of a field, it suffices to show closure under complementation and the formation of finite unions. The proof given here of closure under complementation involves three cases.

Case 1. If $\sum_{k=1}^{n} I_k$ is empty, then $A = I^c$ and $A^c = I$. Since I is either in $\mathcal{B}_I$ or $I = E^1$, in either case I is in the class $\mathcal{B}_o$.

Case 2. If I^c is empty, then $A = \sum_{k=1}^{n} I_k$. Write $I_k = [a_k, b_k)$. Relabel, if necessary, the intervals I_k so that $a_1 \leq b_1 \leq a_2 \leq b_2 \leq \ldots \leq a_n \leq b_n$. That this can be done follows from the fact that A is a disjoint union of intervals. It follows that

$$A^c = [a_1, b_n)^c + \sum_{k=1}^{n-1} [b_k, a_{k+1}) ,$$

so that A^c has the desired form.

Case 3. If neither $\sum_{k=1}^{n} I_k$ nor I^c is empty, then $A^c = I\left(\sum_{k=1}^{n} I_k\right)^c$. But from Case 2

$$\left(\sum_{k=1}^{n} I_k\right)^c = \sum_{k=1}^{n-1} L_k + L^c$$

where L and all the L_k are in $\mathcal{B}_I$. Thus

$$A^c = I\left(\sum_{k=1}^{n-1} L_k + L^c\right) = \sum_{k=1}^{n-1} L_k I + IL^c .$$

Now IL^c is empty, an element of $\mathcal{B}_I$, or a disjoint union of two intervals in $\mathcal{B}_I$. Also, the class $\mathcal{B}_I$ is closed under the formation of finite intersections, as can easily be seen, so that $L_k I \in \mathcal{B}_I$ for each k. Thus A^c is a disjoint union of intervals in $\mathcal{B}_I$ and hence is in $\mathcal{B}_o$. The proof of closure under complementation is now complete.

It remains to verify closure under the formation of finite unions. It suffices to prove that if A and B are in $\mathcal{B}_o$, then $A \cup B \in \mathcal{B}_o$. As before, suppose that $A = \sum_{k=1}^{n} I_k + I^c$ and $B = \sum_{i=1}^{m} J_i + J^c$ be sets in $\mathcal{B}_o$. Two cases arise.

Case 1. Suppose that $(IJ)^c = \emptyset$. This means that $I = J = E^1$. Use the identity $A \cup B = AB^c + AB + A^cB$ to write

$$A \cup B = \left(\sum_{k=1}^{n} I_k\right)\left(\sum_{i=1}^{m} J_i\right)^c + \sum_{k=1}^{n} \sum_{i=1}^{m} I_k J_i + \left(\sum_{k=1}^{n} I_k\right)^c \sum_{i=1}^{m} J_i \ .$$

Let $[a_I, b_I) = \underline{I}$ be the shortest interval half-open on the right which contains $\sum_{k=1}^{n} I_k$. Let $\underline{J}$ be similarly defined for $\sum_{i=1}^{m} J_i$. Then $\left(\sum_{i=1}^{m} J_i\right)^c = \Sigma J_i' + \underline{J}^c$, where the J_i' are intervals in $\underline{J}$ disjoint from the J_i. (In fact, $\Sigma J_i' = \underline{J} - \sum_{i=1}^{m} J_i$.) Similarly $\left(\sum_{k=1}^{n} I_k\right)^c = \Sigma I_k' + \underline{I}^c$.

It follows that

$$A \cup B = \left(\sum_{k=1}^{n} I_k\right)(\Sigma J_i' + \underline{J}^c) + \sum_{i,k} I_k J_i + (\Sigma I_k' + \underline{I}^c) \sum_{i=1}^{m} J_i =$$

$$= \sum_{i,k} I_k J_i' + \sum_{k=1}^{n} I_k \underline{J}^c + \sum_{i,k} I_k J_i + \sum_{i,k} I_k' J_i + \sum_{i=1}^{m} J_i \underline{I}^c \ .$$

Thus $A \cup B$ is a finite disjoint union of intervals in $\mathcal{B}_I$ and therefore is in $\mathcal{B}_o$.

Case 2. Suppose that either I or J is in $\mathcal{B}_I$ so that $IJ = \bar{I} \in \mathcal{B}_I$. Then

$$A \cup B = (A \cup B)E^1 = (A \cup B)(\bar{I} + \bar{I}^c) = (A \cup B)\bar{I} + (A \cup B)\bar{I}^c =$$

$$= \left(\sum_{k=1}^{n} I_k \cup \sum_{i=1}^{m} J_i \cup \bar{I}^c\right)\bar{I} + \left(\sum_{k=1}^{n} I_k \cup \sum_{i=1}^{m} J_i \cup \bar{I}^c\right)\bar{I}^c =$$

$$= \left(\sum_{k=1}^{n} I_k \cup \sum_{i=1}^{m} J_i \cup \bar{I}^c\right)\bar{I} + \bar{I}^c .$$

But $\left(\sum_{k=1}^{n} I_k \cup \sum_{i=1}^{m} J_i \cup \bar{I}^c\right)\bar{I} = \sum_{k=1}^{n} I_k \bar{I} \cup \sum_{i=1}^{m} J_i \bar{I} = (\Sigma I'_k) \cup (\Sigma J'_i)$,

where the I'_k and J'_i are in $\mathcal{B}_I$. This, by Case 1, is reducible to a disjoint union $\sum_{k=1}^{n'} I''_k$ of intervals in $\mathcal{B}_I$. It follows that

$$A \cup B = \sum_{k=1}^{n'} I''_k + \bar{I}^c ,$$

and hence $A \cup B \in \mathcal{B}_o$. This completes the proof. ∎

<u>Theorem 3.4</u>: The Borel field $\mathcal{B}$ is the minimal σ-field over $\mathcal{B}_o$.

<u>Proof</u>: Let $\mathcal{B}^*$ be any σ-field over $\mathcal{B}_o$, and let $\mathcal{B}_J$ be the class of all intervals. Since $\mathcal{B}$ is the minimal σ-field over $\mathcal{B}_J$, it is sufficient to show that $\mathcal{B}^* \supset \mathcal{B}_J$.

Let $I \in \mathcal{B}_J$. If $I = [a, b)$, then by definition $I \in \mathcal{B}_o$. Since

$\mathcal{B}_o \subset \mathcal{B}^*$, $I \varepsilon \mathcal{B}^*$. If $I = [a, b]$, then

$$I = [a, b) \cup \bigcap_{n=1}^{\infty} [b, b + \frac{1}{n}) .$$

Since $\mathcal{B}^*$ is a σ-field and is therefore closed under the formation of countable intersections and countable unions, it follows that $I \varepsilon \mathcal{B}^*$. If $I = (a, b]$, then

$$I = [a, b] - \bigcap_{n=1}^{\infty} [a, a + \frac{1}{n}),$$

and since $\mathcal{B}^*$ is closed under differencing, $I \varepsilon \mathcal{B}^*$. If $I = (a, b)$, then

$$I = (a, b] - \bigcap_{n=1}^{\infty} [b, b + \frac{1}{n})$$

and $I \varepsilon \mathcal{B}^*$. If $I = E^1$ or $I = (-\infty, b)$, then I can clearly be represented as a union of intervals in $\mathcal{B}_I$ and therefore belongs to $\mathcal{B}^*$. Since $[b, \infty) = (-\infty, b)^c$, $[b, \infty) \varepsilon \mathcal{B}^*$, and since $(-\infty, b] = (-\infty, b) + \bigcap_{n=1}^{\infty} [b, b + \frac{1}{n})$, $(-\infty, b] \varepsilon \mathcal{B}^*$. Finally $(b, \infty) = (-\infty, b]^c$ so $(b, \infty) \varepsilon \mathcal{B}^*$. It follows that $B_J \subset \mathcal{B}^*$. This is the desired assertion. ∎

<u>Definition 3.5</u>: Suppose μ is a finite measure on $\mathcal{B}_I$. Define $\bar{\mu}$ on $\mathcal{B}_o$ as follows: $\bar{\mu} A = \sum_{k=1}^{\infty} \mu I_k$ where $\{I_k\}$ is any pairwise disjoint sequence of sets in $\mathcal{B}_I$ for which $\sum_{k=1}^{\infty} I_k = A$, where $A \varepsilon \mathcal{B}_o$.

This definition requires that every $A \varepsilon \mathcal{B}_o$ be representable as an infinite disjoint union of sets in $\mathcal{B}_I$. If A has the form

$\sum_{k=1}^{n} I_k$, then $A = \sum_{k=1}^{\infty} I_k$ where $I_k = \emptyset$ for $k > n$. If A is of the form $\sum_{k=1}^{n} I_k + I^c$, where $I^c \neq \emptyset$ and $I = [a, b)$, then

$$A = \sum_{k=1}^{n} I_k + \sum_{n=o}^{\infty} [a - n - 1, a - n) + \sum_{n=o}^{\infty} [b + n, b + n + 1).$$

In either case A is representable as an infinite disjoint union of sets in $\mathcal{B}_I$. The definition is therefore meaningful.

Lemma 3.6: The relation $\bar{\mu}$ defined above is a function on $\mathcal{B}_o$.

Proof: Suppose that $A \varepsilon \mathcal{B}_o$ has two representations; that is,

$$A = \sum_{k=1}^{n} I_k + I^c = \sum_{r=1}^{m} J_r + J^c .$$

As shown above, I^c can be written as a disjoint union of intervals in $\mathcal{B}_I$, say $I^c = \sum_{k=n+1}^{\infty} I_k$. Thus $\sum_{k=1}^{n} I_k + I^c = \sum_{k=1}^{\infty} I_k$, where all I_k are in $\mathcal{B}_I$. Similarly $\sum_{r=1}^{m} J_r + J^c = \sum_{r=1}^{\infty} J_r$.

$$\bar{\mu}A = \bar{\mu} \quad \sum_{k=1}^{n} I_k + I^c \quad = \sum_{k=1}^{\infty} \mu I_k = \sum_{k=1}^{\infty} \mu I_k A = \sum_{k=1}^{\infty} \mu I_k \quad \sum_{r=1}^{\infty} J_r =$$

$$= \sum_{k=1}^{\infty} \mu \sum_{r=1}^{\infty} I_k J_r = \sum_{k=1}^{\infty} \sum_{r=1}^{\infty} \mu I_k J_r$$

since μ is countably additive on $\mathcal{B}_I$. By interchanging the roles of

$\sum_{k=1}^{n} I_k + I^c$ and $\sum_{r=1}^{m} J_r + J^c$ the same expression can be obtained for $\bar{\mu}\ \sum_{r=1}^{m} J_r + J^c$. Thus $\bar{\mu}$ is a function. ∎

<u>Theorem 3.7</u>: The function $\bar{\mu}$ on $\mathcal{B}_o$ is a measure, and it is the unique extension of μ to $\mathcal{B}_o$.

<u>Proof</u>: To prove that $\bar{\mu}$ is a measure on $\mathcal{B}_o$ it only remains to show that it is countably additive, since it is nonnegative. Assume that $A \varepsilon \mathcal{B}_o$ and $A = \sum_{n=1}^{\infty} A_n$ where each $A_n \varepsilon \mathcal{B}_o$. It must be shown that $\bar{\mu} A = \sum_{n=1}^{\infty} \bar{\mu} A_n$.

Note first that if $B \varepsilon \mathcal{B}_o$ and $B = \sum_{k=1}^{m} J_k + J^c$, where $J^c = \sum_{k=m+1}^{\infty} J_k$ and $J_k \varepsilon \mathcal{B}_I$, then J^c is itself a set in $\mathcal{B}_o$ such that $\bar{\mu} J^c = \sum_{k=m+1}^{\infty} \mu J_k$. Also $\bar{\mu} \sum_{k=1}^{m} J_k = \sum_{k=1}^{m} \mu J_k$. It follows that $\bar{\mu} B = \bar{\mu} \sum_{k=1}^{m} J_k + \bar{\mu} J^c$.

Since A and each A_n are in $\mathcal{B}_o$,

$$A = \sum_{k=1}^{m} I_k + I^c \quad \text{and} \quad A_n = \sum_{k_n=1}^{m_n} I_{k_n}^{(n)}, \quad n = 2,3,\ldots$$

with one set, A_1 say, being $\sum_{k_1=1}^{m_1} I_{k_1}^{(1)} + I^{(1)c}$ where $I^{(1)} \varepsilon \mathcal{B}_I$ or

$I^{(1)} = E^1$.

Clearly either I^c and $I^{(1)^c}$ are either both empty or they are both not. Suppose they are both empty. Then

$$\bar{\mu}A = \sum_{k=1}^{m} \mu I_k = \sum_{k=1}^{m} \mu I_k A = \sum_{k=1}^{m} \mu I_k \sum_{n=1}^{\infty} \sum_{k_n=1}^{m_n} I_{k_n}^{(n)}.$$

Since $I_k \subset A$, $I_k = \sum_{n=1}^{\infty} \sum_{k_n=1}^{m_n} I_k I_{k_n}^{(n)}$. Since μ is countably additive on $\mathcal{B}_I$,

$$\bar{\mu}A = \sum_{k=1}^{m} \sum_{n=1}^{\infty} \sum_{k_n=1}^{m_n} \mu I_k I_{k_n}^{(n)} = \sum_{n=1}^{\infty} \left\{ \sum_{k=1}^{m} \sum_{k_n=1}^{m_n} \mu I_k I_{k_n}^{(n)} \right\} .$$

But

$$A_n = \sum_{k_n=1}^{m_n} I_{k_n}^{(n)} = A \sum_{k_n=1}^{m_n} I_{k_n}^{(n)} = \sum_{k=1}^{m} \sum_{k_n=1}^{m_n} I_k I_{k_n}^{(n)},$$

and each $I_{k_n}^{(n)} I_k \in \mathcal{B}_I$. Hence

$$\bar{\mu}A_n = \sum_{k=1}^{m} \sum_{k_n=1}^{m_n} \mu I_k I_{k_n}^{(n)} .$$

Substitute above to obtain the countable additivity assertion.

Now consider the case where both I^c and $I^{(1)^c}$ are not empty. Since $I^c = I^c I^{(1)^c} + I^c I^{(1)}$,

$$A = \sum_{k=1}^{m} I_k + I^c = \sum_{k=1}^{m} I_k + (I \cup I^{(1)})^c + I^c \ I^{(1)} = \sum_{k=1}^{m'} J_k + J^c,$$

where each $J_k \varepsilon \mathcal{B}_I$ and $J = I \cup I^{(1)}$. Similarly

$$A_1 = \sum_{k_1=1}^{n_1} I_{k_1}^{(1)} + I^{(1)^c} = \sum_{k_1=1}^{n'_1} J_{k_1}^{(1)} + J^c ,$$

and since $I^{(1)^c}$ is disjoint from $A_2, A_3, \ldots, I^{(1)^c} I$ is also disjoint from them. Thus the sets in $\sum_{k_1=1}^{n'_1} J_{k_1}^{(1)}$ maintain the disjoint relationship. Hence it may be assumed that

$$A = \sum_{k=1}^{m} I_k + I^c, \quad \text{and} \quad A_1 = \sum_{k_1=1}^{m_1} I_{k_1}^{(1)} + I^c .$$

Now observe that $\bar{\mu}A = \bar{\mu} \sum_{k=1}^{m} I_k + \bar{\mu} I^c$ and that $\bar{\mu}A_1 = \bar{\mu} \sum_{k_1=1}^{m_1} I_{k_1}^{(1)} + \bar{\mu} I^c$.

Next it is clear that

$$\sum_{k=1}^{m} I_k = \sum_{n=1}^{\infty} \sum_{k_n=1}^{m_n} I_{k_n}^{(n)} .$$

From the preceding case

$$\bar{\mu} \sum_{k=1}^{m} I_k = \bar{\mu} \sum_{k_1=1}^{m_1} I_{k_1}^{(1)} + \sum_{n=2}^{\infty} \bar{\mu} A_n .$$

Therefore

$$\bar{\mu}A = \bar{\mu} \sum_{k=1}^{m} I_k + \bar{\mu} I^c = \bar{\mu} \sum_{k_1=1}^{m_1} I_{k_1}^{(1)} + \bar{\mu} I^c + \sum_{n=2}^{\infty} \bar{\mu} A_n =$$

$$= \bar{\mu} A_1 + \sum_{n=2}^{\infty} \bar{\mu} A_n = \sum_{n=1}^{\infty} \bar{\mu} A_n .$$

This proves that $\bar{\mu}$ on $\mathcal{B}_o$ is a measure.

It remains to show that the extension is unique. So suppose μ' is another extension. Write $A = \sum_{n=1}^{\infty} I_k$, where each $I_k \varepsilon \mathcal{B}_I$ and A is arbitrary in $\mathcal{B}_o$. Then

$$\mu' A = \sum_{n=1}^{\infty} \mu' I_k = \sum_{n=1}^{\infty} \bar{\mu} I_k = \bar{\mu} A .$$

Therefore μ' and $\bar{\mu}$ agree on $\mathcal{B}_o$. The proof is now complete. ∎

<u>Theorem 3.8</u>. The measure μ extends uniquely to $\mathcal{B}_\mu$.

<u>Proof</u>: It has been shown that the Borel field $\mathcal{B}$ is the minimal σ-field over the field $\mathcal{B}_o$. Since μ is assumed finite on $\mathcal{B}_I$, its extension to $\mathcal{B}_o$ is clearly σ-finite on $\mathcal{B}_o$. The Carathéodory extension theorem applies to yield that μ extends uniquely to $\mathcal{B}$.

Let $\bar{\mu}$ denote the extension of μ from $\mathcal{B}_I$ to $\mathcal{B}$, and let $\bar{\bar{\mu}}$ denote the extension of $\bar{\mu}$ from $\mathcal{B}$ to $\mathcal{B}_\mu$. The extension $\bar{\bar{\mu}}$ is defined as follows: if $A = B + N$ where $B \varepsilon \mathcal{B}$ and $N \subset Q \varepsilon \mathcal{B}$ with $\bar{\mu} Q = 0$, then $\bar{\bar{\mu}} A = \bar{\mu} B$. That $\bar{\bar{\mu}}$ is a measure on $\mathcal{B}_\mu$ is obvious.

Suppose μ' is another extension. Then $\mu' A = \mu'(B+N) = \mu' B + \mu' N \leq \mu' B + \mu' Q = \bar{\mu} B + \bar{\mu} Q = \bar{\mu} B = \bar{\bar{\mu}} A$, so that $\mu' A \leq \bar{\bar{\mu}} A$. On the other hand $\mu' A = \mu'(B + N) \geq \mu' B = \bar{\mu} B = \bar{\bar{\mu}} A$ so that $\mu' A \geq \bar{\bar{\mu}} A$. It follows that μ' and $\bar{\bar{\mu}}$ agree on $\mathcal{B}_\mu$. This completes the proof. ∎

Theorems 3.3, 3.4, 3.7, and 3.8 give the results stated in Part I.

The Proof of Part II

In order for the relation $\sim$ to be an equivalence relation it is necessary and sufficient that $F \sim G \Longrightarrow G \sim F$; $F \sim G$ and $G \sim H \Longrightarrow F \sim H$; and $F \sim F$; for all F, G, and H in $\mathfrak{F}$. It is easily seen that these three conditions hold, so that $\sim$ is an equivalence relation. It is also clear that μ is finite. To verify the remaining assertion it must be shown that if $I = \sum_{n=1}^{\infty} I_n$, where $I = [a, b)$ and $I_n = [a_n, b_n)$, then $\mu I = \sum_{n=1}^{\infty} \mu I_n$.

Let n be given. Then each $I_j \subset I$, $j = 1,2,\ldots,n$, and the I_j are pairwise disjoint. Interchange the subscripts, if necessary, so that $a \leq a_1 \leq b_1 \leq a_2 \leq b_2 \leq \cdots \leq a_n \leq b_n \leq b$. Then

$$\sum_{k=1}^{n} \mu I_k = \sum_{k=1}^{n} \Delta F(a_k; b_k - a_k) \leq \sum_{k=1}^{n} [F(b_k) - F(a_k)] + $$
$$+ \sum_{k=1}^{n+1} [F(a_k) - F(b_{k-1})]$$

since the last written sum is nonnegative. Here $F(b_o) = F(a)$ and $F(a_{n+1}) = F(b)$. Thus

$$\sum_{k=1}^{n} \mu I_k \leq \Big\{[F(b_1) - F(a_1)] + [F(b_2) - F(a_2)] + \ldots + [F(b_n) - F(a_n)]\Big\} +$$
$$+ \Big\{[F(a_1) - F(a)] + [F(a_2) - F(b_1)] + \ldots + [F(a_n) - F(b_{n-1})] +$$
$$+ [F(b) - F(b_n)]\Big\} = F(b) - F(a) = \Delta F(a; b - a) = \mu I .$$

Since this holds for all n,

$$\sum_{n=1}^{\infty} \mu I_n \leq \mu I .$$

It remains to prove the reverse inequality. If $b = a$, there is nothing to prove; so assume that $a < b$. Let $\varepsilon > 0$ be arbitrary. Set $I^\delta = [a, b - \delta)$ where, using left-continuity of F, δ is chosen so that $a < b - \delta$ and $\Delta F(b - \delta; \delta) < \frac{\varepsilon}{2}$. Also due to left-continuity, for every n there is a $\delta_n > 0$ such that

$$\Delta F(a_n - \delta_n; \delta_n) < \frac{\varepsilon}{2^{n+1}} .$$

Write $I_n^\delta = (a_n - \delta_n, b_n)$. Then, since $I^\delta \subset \bigcup_{n=1}^{\infty} I_n^\delta$, it follows from the Heine-Borel theorem that some finite subclass of the I_n^δ also covers I^δ. It is assumed that no interval in this subclass is redundant; for if that be the case, the redundant intervals may be eliminated. Thus

$$I^\delta \subset \bigcup_{k=1}^{m} I_{n_k}^\delta .$$

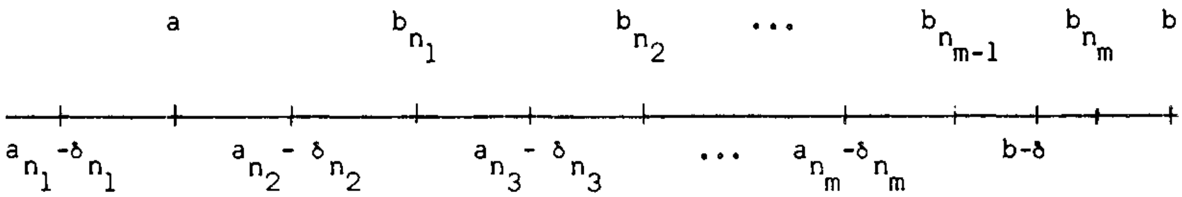
Also there is no loss of generality in assuming the subscripts are ordered so that

$$a_{n_1} - \delta_{n_1} < a;\ b_{n_m} > b - \delta;\ a_{n_1} - \delta_{n_1} < a_{n_2} - \delta_{n_2};\ b_{n_{m-1}} < b_{n_m}$$

and so that for $i = 2,3,\ldots,m-1$,

$$a_{n_i} - \delta_{n_i} < b_{n_{i-1}} < a_{n_{i+1}} - \delta_{n_{i+1}} < b_{n_i} .$$

The situation is illustrated below.



Now

$$\Delta F(a;\, b - a - \delta) \leq \Delta F(a_{n_1} - \delta_{n_1};\, b_{n_m} - a_{n_1} + \delta_{n_1}) =$$

$$= \Delta F(a_{n_1} - \delta_{n_1};\, b_{n_1} - a_{n_1} + \delta_{n_1}) + \sum_{k=2}^{m} \Delta F(b_{n_{k-1}};\, b_{n_k} - b_{n_{k-1}}) .$$

Since $\Delta F(b_{n_{k-1}};\, b_{n_k} - b_{n_{k-1}}) \leq \Delta F(a_{n_k} - \delta_{n_k};\, b_{n_k} - a_{n_k} + \delta_{n_k})$, if the former is replaced by the latter in the last written sum, the inequality becomes

$$\Delta F(a;\, b - a - \delta) \leq \sum_{k=1}^{m} \Delta F(a_{n_k} - \delta_{n_k};\, b_{n_k} - a_{n_k} + \delta_{n_k}) .$$

But

$$\sum_{k=1}^{m} \Delta F(a_{n_k} - \delta_{n_k};\, b_{n_k} - a_{n_k} + \delta_{n_k}) \leq \sum_{n=1}^{\infty} \Delta F(a_n - \delta_n;\, b_n - a_n + \delta_n) =$$

$$= \sum_{n=1}^{\infty} \left\{ \Delta F(a_n;\, b_n - a_n) + \Delta F(a_n - \delta_n;\, \delta_n) \right\} = \sum_{n=1}^{\infty} \Delta F(a_n;\, b_n - a_n) +$$

$$+ \sum_{n=1}^{\infty} \Delta F(a_n - \delta_n;\, \delta_n) \leq \sum_{n=1}^{\infty} \Delta F(a_n;\, b_n - a_n) + \frac{\varepsilon}{2} .$$

The above inequalities then reduce to

$$\Delta F(a;\ b - a - \delta) \leq \sum_{n=1}^{\infty} \Delta F(a_n;\ b_n - a_n) + \frac{\varepsilon}{2}\ .$$

Finally

$$\Delta F(a;\ b - a) = \Delta F(a;\ b - a - \delta) + \Delta F(b - \delta;\ \delta) <$$

$$< \Delta F(a;\ b - a - \delta) + \frac{\varepsilon}{2} < \sum_{n=1}^{\infty} \Delta F(a_n;\ b_n - a_n) + \varepsilon\ .$$

Since this holds for every $\varepsilon > 0$,

$$\Delta F(a;\ b - a) \leq \sum_{n=1}^{\infty} \Delta F(a_n;\ b_n - a_n)\ .$$

This is the reverse inequality. Therefore μ on $\mathcal{B}_I$ is countably additive and is hence a measure. This completes the proof of Part II. ∎

The Proof of Part III

Let $\mathcal{M}(\mathcal{B}_I)$ be the class of Lebesgue-Stieltjes measures on $\mathcal{B}_I$, and let φ be the function which assigns to each equivalence class $\mathcal{F}_F \varepsilon \mathcal{F}$ the measure $\mu[a, b) = \Delta F(a;\ b - a)$. By Part II, $\varphi(\mathcal{F}) \subset \mathcal{M}(\mathcal{B}_I)$. If $\mu \varepsilon \mathcal{M}(\mathcal{B}_I)$ and $\mu = \varphi(\mathcal{F}_F) = \varphi(\mathcal{F}_G)$, then $\Delta F = \Delta G$ by definition of φ, and hence $\mathcal{F}_F = \mathcal{F}_G$. It is therefore sufficient to show that φ maps $\mathcal{F}$ onto $\mathcal{M}(\mathcal{B}_I)$. To do so, a distribution function F will be constructed so that $\Delta F = \mu$ for any given $\mu \varepsilon \mathcal{M}(\mathcal{B}_I)$. The function F defined by

$$F(x) = \begin{cases} \mu[o,\ x), & \text{if } x \geq o\ ; \\ -\mu[x,\ o), & \text{if } x \leq o \end{cases}$$

is such a distribution function. To verify that $\Delta F = \mu$, three cases will be considered.

Case 1. If $0 \leq a_o \leq a_1$, then from the additivity of μ on $\mathcal{B}_I$, $\mu[a_o, a_1) = \mu[o, a_1) - \mu[o, a_o)$, or $\mu[a_o, a_1) = F(a_1) - F(a_o) =$ $= \Delta F(a_o; a_1 - a_o)$.

Case 2. If $a_o \leq 0 \leq a_1$, then $\mu[a_o, a_1) = \mu[a_o, 0) + \mu[0, a_1)$, or $\mu[a_o, a_1) = - F(a_o) + F(a_1) = \Delta F(a_o; a_1 - a_o)$.

Case 3. If $a_o \leq a_1 \leq 0$, then $\mu[a_o, a_1) = \mu[a_o, 0) - \mu[a_1, 0)$, or $\mu[a_o, a_1) = - F(a_o) + F(a_1) = \Delta F(a_o, a_1 - a_o)$.

Thus $\Delta F = \mu$.

It remains to show that F is a distribution function. Clearly F is finite. If $x_1 \leq x_2$, then since $\mu = \Delta F$, $F(x_2) - F(x_1) =$ $= \mu[x_1, x_2) \geq 0$ so that F is nondecreasing. It follows from the continuity theorem for measures that

$$\lim_{x_1 \to x_2^-} \{F(x_2) - F(x_1)\} = \lim_{x_1 \to x_2^-} \mu[x_1, x_2) = \mu \lim_{x_1 \to x_2^-} [x_1, x_2) = \mu\emptyset = 0,$$

so that F is continuous from the left. F is therefore a distribution function and Part III is proved.

This completes the proof of the theorem in E^1.

CHAPTER IV

THE CORRESPONDENCE THEOREM IN N DIMENSIONS

In this chapter the correspondence theorem in E^N is proved. The proof mimics the proof of the theorem in E^1.

The Proof of Part I

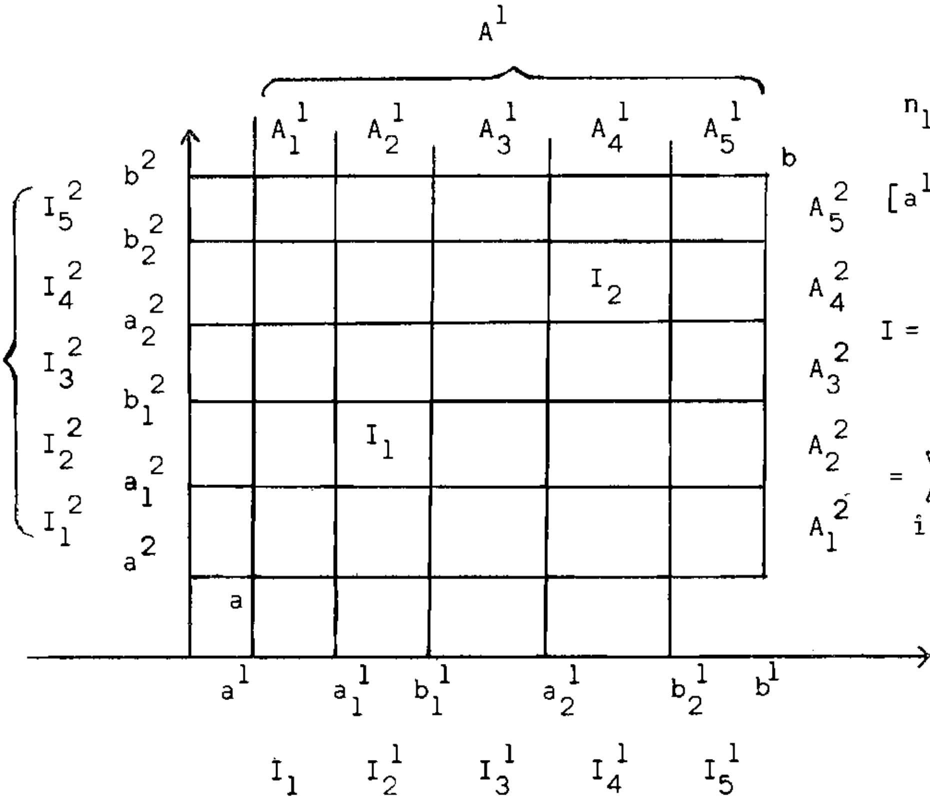
The definitions and theorems of Part I remain the same in E^N (with obvious modifications) except for the following.

Distribution functions F on E^N are described just as those on E^1 except that the nondecreasing condition is replaced by the monotone requirement $\Delta_{1\ldots N} F \geq 0$ given in Theorem 2.10.

The ordering process used to prove that $\mathcal{B}_o$ is a field in the one-dimensional case is more complicated in the multidimensional case. It is described in the following

Theorem 4.1: The class $\mathcal{B}_o$ is a field.

Proof: The ordering process for the case $N = n = 2$ is illustrated below.



$n_1 = n_2 = 5.$

$[a^1, b^1) = \sum_{i=1}^{5} I_k^1.$

$I = \bigcap_{i=1}^{2} A^i = \bigcap_{i=1}^{2} \sum_{j=1}^{5} A_i^j =$

$= \sum_{i,j} A_i^1 A_j^2 = \sum_{i,j} I_i^1 I_j^2.$

It will first be shown that if $I \supset \sum_{k=1}^{n} I_k$ where I and the I_k are in $\mathcal{B}_I$, then $I - \sum_{k=1}^{n} I_k$ is a disjoint sum of intervals in $\mathcal{B}_I$. Write $I = [a, b) = \bigtimes_1^N [a^t, b^t) = \bigcap_1^N A^t$, where $A^t = E_1 \times \dots \times [a^t, b^t) \times \times \dots \times E_N$. Similarly write $I_k = \bigtimes_1^N [a_k^t, b_k^t) = \bigcap_1^N A_k^t$. Now, for each t, rearrange the a_k^t and b_k^t in ascending order, and let them be relabeled as follows:

$$a^t = x_o^t \leq x_1^t \leq \dots \leq x_{n_t}^t = b^t .$$

Denote $[x_{k_t-1}^t, x_{k_t}^t)$ by $I_{k_t}^t$ and $E_1 \times \dots \times I_{k_t}^t \times \dots \times E_N$ by $A_{k_t}^t$,

$k_t = 1,2,\ldots,n_t$, $t = 1,2,\ldots, N$. Then

$$[a^t, b^t) = \sum_{k_t=1}^{n_t} I^t_{k_t}, \quad A^t = \sum_{k_t=1}^{n_t} A^t_{k_t}, \quad \text{and} \quad I = \bigcap_{t=1}^{N} \sum_{k_t=1}^{n_t} A^t_{k_t}.$$

By a distributive law the last intersection can be written as a sum of intersections of the $A^t_{k_t}$. Also, each of these intersections of the $A^t_{k_t}$ can be written as a cross product of the $I^t_{k_t}$. Thus I can be written

$$I = \sum_{k_t=1}^{n'} \bigtimes_{t=1}^{N} I^t_{k_t}.$$

Some of these basic intervals $\bigtimes_{t=1}^{N} I^t_{k_t}$ add to the intervals I_k, while the sum of the remaining ones is $I - \sum_{1}^{N} I_k = \sum_{1}^{m} I'_k$, say, where the I'_k are intervals in $\mathcal{B}_I$.

Let $A = \sum_{k=1}^{n} I_k + I^c$ and $B = \sum_{k=1}^{m} J_k + J^c$ be arbitrary sets in $\mathcal{B}_o$. As in the proof in E^1 there are three cases considered in proving that $\mathcal{B}_o$ is closed under complementation.

Case 1. If $\sum_{k=1}^{n} I_k$ is empty, then $A = I^c$ and $A^c = I$. I is in $\mathcal{B}_o$ since I is either in $\mathcal{B}_I$ or $I = E^N$.

Case 2. If I^c is empty, then $A = \sum_{k=1}^{n} I_k$. There is an interval $\bar{I} \supset \sum_{k=1}^{n} I_k$. It has been shown that

$$\bar{I} - \sum_{k=1}^{n} I_k = \sum_{1}^{m} J_r$$

where the J_r are in $\mathcal{B}_I$. Thus

$$\left(\sum_{k=1}^{n} I_k\right)^c = \left(\bar{I} - \sum_{r=1}^{m} J_r\right)^c = \sum_{r=1}^{m} J_r + \bar{I}^c .$$

Hence $\left(\sum_{k=1}^{n} I_k\right)^c \varepsilon \mathcal{B}_o$.

Case 3. If neither $\sum_{k=1}^{n} I_k$ nor I^c is empty, then $A^c = \bigcap_{k=1}^{n} I\, I_k^c$. Since $I \supset I_k$, $k = 1,\ldots,n$, and since $I\, I_k^c = I - I_k$,

$$I\, I_k^c = \sum_{r_k=1}^{m_k} J_{r_k} ,$$

where the $J_{r_k} \varepsilon \mathcal{B}_I$. Thus

$$A^c = \bigcap_{k=1}^{n} \sum_{r_k=1}^{m_k} J_{r_k} .$$

By the distributive law this intersection can be written as a sum of intersections of the J_{r_k}. It can easily be shown that the class $\mathcal{B}_I$ is closed under the formation of finite intersections. It follows that A^c can be written as a sum of intervals in $\mathcal{B}_I$ and hence $A^c \varepsilon \mathcal{B}_o$. This completes the proof of closure under complementation.

The proof of closure under the formation of finite unions follows now just as in the one-dimensional case.

The proof is now complete. ∎

<u>Theorem 4.2</u>. $\mathcal{B}$ is the minimal σ-field over $\mathcal{B}_o$.

Proof: As in the one-dimensional case it is sufficient to show that any σ-field $\mathcal{B}^*$ over $\mathcal{B}_o$ contains the class of all intervals. So let $I = \bigtimes_{t=1}^{N} I^t$ be any interval at all. Write $I = \bigcap_{t=1}^{N} A^t$, where $A^t = E_1 \times \ldots \times I^t \times \ldots \times E_N$, and for $t' \neq t$ let $V_n^{t'} = [-n, n)$, $n = 1,2,\ldots$. Let $\{I_m^t\}$ be a sequence of one-dimensional intervals of $\mathcal{B}_I$ for which $\bigcap_1^\infty I_m^t = I^t$ or $\bigcup_1^\infty I_m^t = I^t$. (Such a sequence can always be constructed.) Clearly, either

$$A^t = \bigcup_{m=1}^\infty \bigcup_{n=1}^\infty (V_n^1 \times \ldots \times I_m^t \times \ldots \times V_n^N) \text{ or}$$

$$A^t = \bigcap_{m=1}^\infty \bigcup_{n=1}^\infty (V_n^1 \times \ldots \times I_m^t \times \ldots \times V_n^N),$$

so that either

$$I = \bigcap_{t=1}^N \bigcup_{m=1}^\infty \bigcup_{n=1}^\infty (V_n^1 \times \ldots \times I_m^t \times \ldots \times V_n^N) \text{ or}$$

$$I = \bigcap_{t=1}^N \bigcap_{m=1}^\infty \bigcup_{n=1}^\infty (V_n^1 \times \ldots \times I_m^t \times \ldots \times V_n^N).$$

In either case $I \in \mathcal{B}^*$ since $\mathcal{B}^*$ is a σ-field over $\mathcal{B}_o$ and is therefore closed under the formation of countable unions and countable intersections. This completes the proof. ∎

The only remaining modification in Part I is the verification that any set in $\mathcal{B}_o$ can be represented as an infinite sum of intervals in $\mathcal{B}_I$. It suffices to show that if $I \in \mathcal{B}_I$, where $I \neq \emptyset$, then $I^c = \sum_{n=1}^\infty I_n$, where each $I_n \in \mathcal{B}_I$. Clearly there is a sequence $\{J_n\}$ of intervals in

$\mathcal{B}_I$ such that $J_1 = I$, $J_n \subset J_{n+1}$, $n = 1,2,\ldots,$ and $\bigcup_{n=1}^{\infty} J_n = E^N$. Also it is clear that $I^c = \sum_{n=1}^{\infty} (J_{n+1} - J_n)$. By the first result in Theorem 4.1 $J_{n+1} - J_n = \sum_{k=1}^{m_n} I_{nk}$, where $I_{nk} \in \mathcal{B}_I$, $k = 1,\ldots,m_n$; $n = 1,2,\ldots$. Thus

$$I^c = \sum_{n=1}^{\infty} \sum_{k=1}^{m_n} I_{nk}$$

so that I^c has the desired form.

Difficulties arise in generalizing the proofs in Part II. That the relation, $\sim$, on the class $\mathcal{F}$ of distribution functions in E^N, defined by

$$F \sim G \text{ if and only if } \Delta_{12\ldots N}F = \Delta_{12\ldots N}G$$

for all $x \in E^N$, is an equivalence relation and that the set function $\mu = \Delta_{1\ldots N}F$ is finite and nonnegative are obvious assertions. It remains to prove that μ is σ-additive on $\mathcal{B}_I$. This is the difficult part.

The proof of σ-additivity will require finite additivity, and the proof of finite additivity requires several preliminaries to which the following section is devoted.

Preliminaries to Generalizing Proof of Part II

Definition 4.3: Let $A \subset E^N$, $T_N = \{1,2,\ldots,N\}$, and let $T_M = \{t_1, t_2,\ldots,t_M\}$ be a subset of T_N. The projection A_{T_M} of A onto $E^{T_M} = \bigtimes_{j=1}^{M} E_{t_j}$ is defined as follows:

$$A_{T_M} = \left\{ x_{T_M} \varepsilon\ E^{T_M} \;\middle|\; \left\{x_{T_M}\right\} \times \left\{x_{T_{N-M}}\right\} \subset A \text{ for some } x_{T_N - T_M} \varepsilon\ E^{T_N - T_M} \right\} .$$

Here we have written $E^{T_N - T_M} = \bigtimes_{t\,\varepsilon\, T_N - T_M} E_t$.

Corollary 4.4: $A_{T_M} \cup B_{T_M} = (A \cup B)_{T_M}$; $\left(\bigtimes_{t=1}^{N} A_t \right)_{T_M} = \bigtimes_{t \varepsilon T_M} A_t$,

where each $A^t \varepsilon\ E^t$; and if $T_R \subset T_M$, $(A_{T_M})_{T_R} = A_{T_R}$.

Proof: The proof follows from set inclusion arguments. The details will not be given here.

Theorem 4.5: Let $I_1 = \bigtimes_{t=1}^{N} I_1^t$ and $I_2 = \bigtimes_{t=1}^{N} I_2^t$, where for $j = 1,2,$ $I_j^t = [a_j^t , b_j^t)$ is a one-dimensional interval. A necessary and sufficient condition for $I_1 + I_2 = I = \bigtimes_{t=1}^{N} I^t$, where $I^t = [a^t, b^t)$ is a one-dimensional interval, is that for all t save one $I_1^t = I_2^t$, and that for the remaining ones the one-dimensional intervals abut.

Proof: The sufficiency of the condition will be proved first. Relabel, if necessary, so that I_1^1 and I_2^1 abut. That is, $I_1^1 + I_2^1 = [a_1^1 , b_1^1) + [a_2^1 , b_2^1) = [a_1^1 , b_2^1)$. Now use the device $\bigtimes_{t \varepsilon T_N} I_j^t = \bigcap_{t \varepsilon T_N} A_j^t$ where $A_j^t = E_1 \times \ldots \times I_j^t \times \ldots \times E_N$, $t = 1,2,\ldots,N$, $j = 1,2,$ with $T_N = \{1,2,\ldots,N\}$.
It follows that $I_1 = \bigcap_{t=1}^{N} A_1^t$, $I_2 = \bigcap_{t=1}^{N} A_2^t$, that $A_1^1 + A_2^1 =$
$= (I_1^1 + I_2^1) \times E_2 \times \ldots \times E_N$ and that $A_1^t = A_2^t$ for $t = 2,\ldots,N$. Clearly I_1 and I_2 are disjoint and

$$I_1 + I_2 = \bigcap_{t=1}^{N} A_1^t + \bigcap_{t=1}^{N} A_2^t = \left(\bigcap_{t=2}^{N} A^t\right)\left(A_1^1 + A_2^1\right) =$$

$$= \left(\bigtimes_{t=2}^{N} I_1^t\right) \times \left(I_1^1 + I_2^1\right)$$

so that $I_1 + I_2 = I$. The condition is therefore sufficient.

Conversely suppose that $I_1 + I_2 = I$. It is easily seen that for some t, $I_1^t I_2^t$ is empty. For if not, there is an x_t in $I_1^t I_2^t$ for each $t = 1,2,\ldots,N$ so that $(x_1, x_2,\ldots,x_N) \in I_1 I_2$ since $I_1 I_2 = \bigtimes_{t=1}^{N} (I_1^t I_2^t)$; and therefore I_1 is not disjoint from I_2. Relabel, if necessary, so that I_1^1 and I_2^1 are disjoint. It will now be shown that I_1^1 and I_2^1 abut. For definiteness let I_1^1 be to the left of I_2^1. Suppose they do not abut. Then $b_1^1 < a_2^1$ and $I_1^1 + I_2^1$ is not an interval. But $I_1 + I_2$ is an interval (in E^N).

By Corollary 4.4 the projection of $I_1 + I_2$ onto E_1 (the projection is an interval in E_1) must be equal to the union of the projections of I_1 and I_2. Thus $I_1^1 + I_2^1$ is an interval. Hence we must reject the supposition that I_1^1 and I_2^1 do not abut.

It remains to prove that $I_1^t = I_2^t$ for all remaining t. Suppose this is not the case. Relabel the intervals so that $I_1^2 \neq I_2^2$. Corollary 4.4 and the assumptions that I_1, I_2, and $I_1 + I_2$ are all intervals yield that the projection of I onto E_1 is $I_1^1 + I_2^1$, that the projection of I onto E_2 is the union of the projection of I_1 onto E_2 and the projection of I_2 onto E_2, and that this union is $I_1^2 \cup I_2^2$. Again by the corollary the projection of I onto $E_1 \times E_2$ is $(I_1^1 \times I_1^2) + (I_2^1 \times I_2^2)$ and is also an interval. Moreover the projections of $(I_1^1 \times I_1^2) + (I_2^1 \times I_2^2)$ onto E_1 and E_2 are, respectively,

$I_1^1 + I_2^1$ and $I_1^2 \cup I_2^2$. The projections of I onto E_1, E_2, and $E_1 \times E_2$ are also I^1, I^2 and $I^1 \times I^2$. By the above, then, $I^1 = I_1^1 + I_2^1$, $I^2 = I_1^2 \cup I_2^2$ and $I^1 \times I^2 = (I_1^1 \cup I_1^2) + (I_2^1 \times I_2^2)$. It follows that

$$(I_1^1 \times I_1^2) + (I_2^1 \times I_2^2) = (I_1^1 + I_2^1) \times (I_1^2 \cup I_2^2).$$

Now suppose that $b \in I_2^2$. If $b \notin I_1^2$, then for any $a \in I_1^1$, $(a, b) \in (I_1^1 + I_2^1) \times (I_1^2 \cup I_2^2)$. But since $b \notin I_1^2$, $(a, b) \notin I_1^1 \times I_1^2$; and since I_1^1 and I_2^1 are disjoint, $a \notin I_2^1$ so that $(a, b) \notin I_2^1 \times I_2^2$. This contradicts the last equality. Hence $b \in I_1^2$ and $I_2^2 \subset I_1^2$. The reverse inclusion is proved similarly. Therefore the last equality implies that $I_1^2 = I_2^2$. This proves the theorem. ∎

<u>Theorem 4.6</u>: If μ is defined on $\mathcal{B}_I$ by the rule, $\mu[a,b) = \Delta_{1\ldots N}F(a;b-a)$, where F is a distribution function and $-\infty < a \leq b < \infty$, and if I_1, I_2, and $I_1 + I_2 = I$ are all intervals in $\mathcal{B}_I$, then $\mu I = \mu I_1 + \mu I_2$.

<u>Proof</u>: By Theorem 4.5, I_1, I_2, and I can be written

$$I_1 = [a_1^1, b_1^1) \times \bigtimes_{t=2}^{N} [a^t, b^t), \quad I_2 = [a_2^1, b_2^1) \times \bigtimes_{t=2}^{N} [a^t, b^t),$$

and
$$I = [a_1^1, b_2^1) \times \bigtimes_{t=2}^{N} [a^t, b^t)$$

where $b_1^1 = a_2^1$. By Theorem 2.2 and Definition 2.1

$$\mu I_1 = \Delta_1[\Delta_{2\ldots N}F(a_1^{\,1}, a^2,\ldots,a^N;\ b^2-a^2,\ldots,b^N-a^N);\ b_1^1-a_1^1] =$$

$$= \Delta_{2\ldots N}\, F(b_1^{\,1}, a^2,\ldots,a^N;\ b^2-a^2,\ldots,b^N-a^N) +$$

$$- \Delta_{2\ldots N}\, F(a_1^{\,1}, a^2,\ldots,a^N;\ b^2-a^2,\ldots,b^N-a^N).$$

Similarly

$$\mu I_2 = \Delta_{2\ldots N}\, F(b_2^{\,1}, a^2,\ldots,a^N;\ b^2-a^2,\ldots,b^N-a^N) +$$

$$- \Delta_{2\ldots N}\, F(a_1^{\,1}, a^2,\ldots,a^N;\ b^2-a^2,\ldots,b^N-a^N)$$

so that

$$\mu I_1 + \mu I_2 = \Delta_{2\ldots N}\, F(b_2^{\,1}, a^2,\ldots,a^N;\ b^2-a^2,\ldots,b^N-a^N) +$$

$$- \Delta_{2\ldots N}\, F(a_1^{\,1}, a^2,\ldots,a^N;\ b^2-a^2,\ldots,b^N-a^N) =$$

$$= \Delta_{12\ldots N}\, F(a_1^{\,1}, a^2,\ldots,a^N;\ b_2^1-a_1^1,\ b^2-a^2,\ldots,b^N-a^N) = \mu I.$$

This completes the proof. ∎

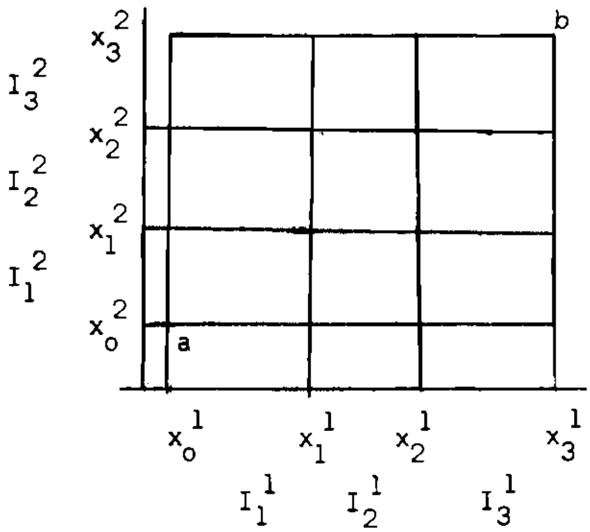
<u>Theorem 4.7</u>: Let $I = [a, b) = \bigtimes_{t=1}^{N} [a^t, b^t) = \bigtimes_{t=1}^{N} I^t$ be an interval in $\mathcal{B}_I$. For $t = 1,2,\ldots,N$ form partitions of I^t: $a^t = x_0^t \leq x_1^t \leq \ldots \leq x_{n_t}^t = b^t$. Denote $[x_{k_t-1}^t, x_{k_t}^t)$ by $I_{k_t}^t$, $k_t = 1,2,\ldots,n_t$. Then

$$[a, b) = \sum_{k_N=1}^{n_N} \ldots \sum_{k_2=1}^{n_2} \sum_{k_1=1}^{n_1} \bigtimes_{t=1}^{N} I_{k_t}^t,$$

and for the μ of the previous theorem

$$\mu[a,\ b) = \sum_{k_N=1}^{n_N} \cdots \sum_{k_2=1}^{n_2} \sum_{k_1=1}^{n_1} \mu \bigtimes_{t=1}^{N} I_{k_t}^t .$$

<u>Proof</u>: The situation in E^2 is illustrated below.



$$[a,b) = \sum_{k_2=1}^{3} \sum_{k_1=1}^{3} (I_{k_1}^1 \times I_{k_2}^2) .$$

It is easily seen that $I^t = \sum_{k_t=1}^{n_t} I_{k_t}^t$ $(t = 1,2,\ldots,N)$. Let A^t and $A_{k_t}^t$ be cylinders constructed, respectively, on I^t and $I_{k_t}^t$. That is,

$$A^t = I^t \times (\bigtimes_{t' \neq t} E_{t'}) \ , \quad A_{k_t} = I_{k_t}^t \times (\bigtimes_{t' \neq t} E_{t'}) .$$

Then $A^t = \sum_{k_t=1}^{n_t} A_{k_t}^t$ and

$$[a,\ b) = \bigtimes_{t=1}^{N} I^t = \bigcap_{t=1}^{N} A^t = \bigcap_{t=1}^{N} \sum_{k_t=1}^{n_t} A_{k_t}^t .$$

By the distributive law, the last intersection can be written

$$\sum_{k_1=1}^{n_1} \cdots \sum_{k_N=1}^{n_N} \bigcap_{t=1}^{N} A^t_{k_t} = \sum_{k_1=1}^{n_1} \cdots \sum_{k_N=1}^{n_N} \bigtimes_{t=1}^{N} I^t_{k_t} .$$

It remains to prove the measure additivity statement. Since

$$\sum_{k_1=1}^{n_1} \left(\bigtimes_{t=1}^{N} I^t_{k_t} \right) = \sum_{k_1=1}^{n_1} \left[I^1_{k_1} \times \left(\bigtimes_{t=2}^{N} I^t_{k_t} \right) \right],$$

the fact that $I^1_{k_1}$ and $I^1_{k_1+1}$ are adjoint and $\sum_{k_1=1}^{n_1} I^1_{k_1} = I^1$ guarantees that for any integer j, $1 \leq j \leq n_1$, the sum $\sum_{k_1=1}^{j} \bigtimes_{t=1}^{N} I^t_{k_t}$ is an interval. So, for fixed $k_2, k_3, \ldots, k_N$,

$$\sum_{k_1=1}^{n_1} \mu \bigtimes_{t=1}^{N} I^t_{k_t} = \mu \sum_{k_1=1}^{n_1} \bigtimes_{t=1}^{N} I^t_{k_t}$$

by a repeated application of the previous theorem. The last expression is $\mu\, I^1 \times (\bigtimes_{t=2}^{N} I^t_{k_t})$. Similarly

$$\sum_{k_2=1}^{n_2} \sum_{k_1=1}^{n_1} \mu \bigtimes_{t=1}^{N} I^t_{k_t} = \sum_{k_2=1}^{n_2} \mu\, I^1 \times (\bigtimes_{t=2}^{N} I^t_{k_t}) =$$

$$= \mu \sum_{k_2=1}^{n_2} I^1 \times (\bigtimes_{t=2}^{N} I^t_{k_t}) = \mu\, I^1 \times I^2 \times (\bigtimes_{t=3}^{N} I^t_{k_t}).$$

The conclusion follows by an obvious induction. ∎

The last two theorems illustrate the problem of generalizing the correspondence theorem: disjoint intervals which do not pairwise add up to intervals can block together to form an interval.

We now return to finite additivity and σ-additivity.

The Proof of Finite Additivity and σ-Additivity

Theorem 4.8: The set function $\mu = \Delta_{1\ldots N}F$, where F is a distribution function, is finitely additive on $\mathcal{B}_I$.

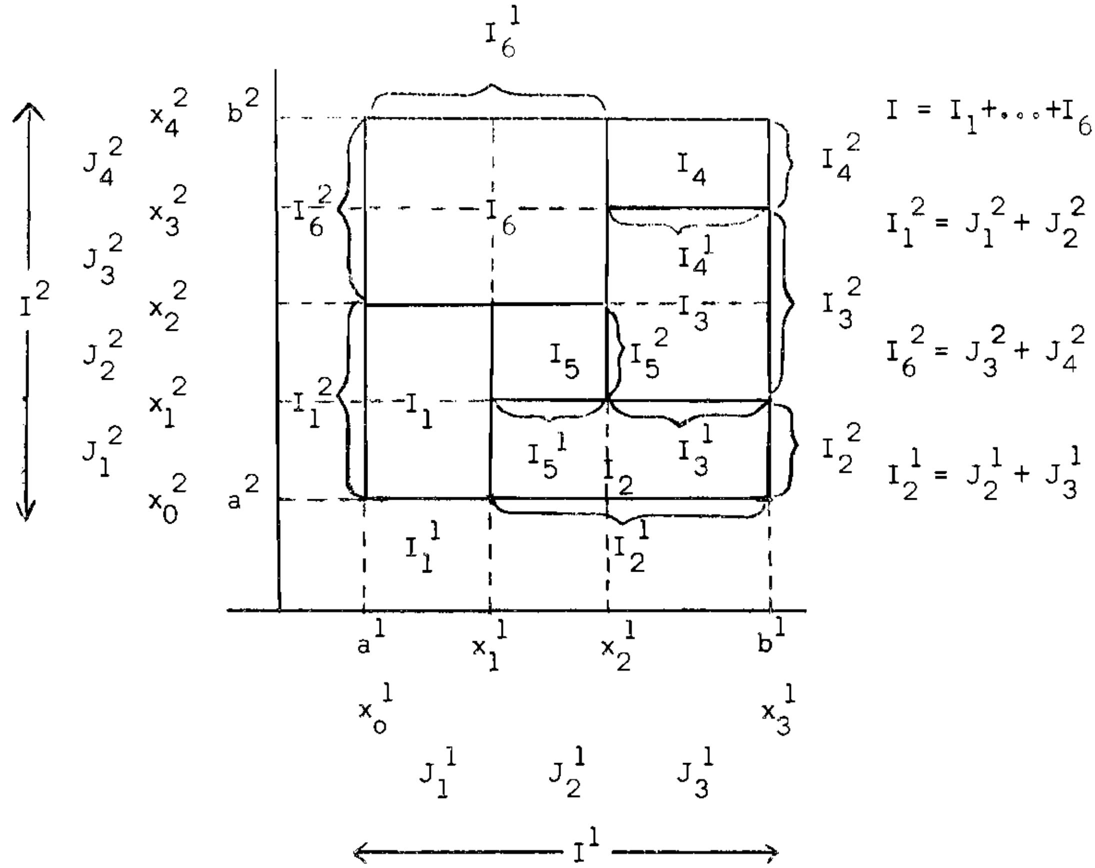
Proof: Let $I = \sum_{r=1}^{m} I_r$, where $I = \bigtimes_{t=1}^{N} [a^t, b^t) = \bigtimes_{t=1}^{N} I^t$ and $I_r = \bigtimes_{t=1}^{N} [a_r^t, b_r^t) = \bigtimes_{t=1}^{N} I_r^t$. For fixed t consider all the numbers a_r^t and b_r^t $(r = 1,2,\ldots, m)$. Order these into $x_o^t \leq x_1^t \leq \cdots \leq x_{n_t}^t$. Write $J_{k_t}^t = [x_{k_t-1}^t, x_{k_t}^t)$. Then

$$I^t = [a^t, b^t) = \sum_{k_t=1}^{n_t} J_{k_t}^t .$$

Also, for given t and r, there is an $x_{\alpha_{tr}}^t = a_r^t$ and an $x_{\beta_{tr}}^t = b_r^t$, so that

$$I_r^t = [a_r^t, b_r^t) = \sum_{k_t=\alpha_{tr}}^{\beta_{tr}} J_{k_t}^t ,$$

where $J_{k_t}^t \in \mathcal{B}_I$. The situation in E^2 is illustrated below.



By analogy with the previous theorem

$$I_r = \underset{t=1}{\overset{N}{\times}} I_r^t = \sum_{k_1=\alpha_{1r}}^{\beta_{1r}} \cdots \sum_{k_N=\alpha_{Nr}}^{\beta_{Nr}} \underset{t=1}{\overset{N}{\times}} J_{k_t}^t .$$

Since $I = \sum_{r=1}^{m} I_r$,

$$\sum_{r=1}^{m} \left[\sum_{k_1=\alpha_{1r}}^{\beta_{1r}} \cdots \sum_{k_N=\alpha_{Nr}}^{\beta_{Nr}} \underset{t=1}{\overset{N}{\times}} J_{k_t}^t \right] = \sum_{k_1=1}^{n_1} \cdots \sum_{k_N=1}^{n_N} \underset{t=1}{\overset{N}{\times}} J_{k_t}^t ,$$

each side being a grouping of the same intervals. Also, by the previous theorem,

$$\mu I = \sum_{k_1=1}^{n_1} \cdots \sum_{k_N=1}^{n_N} \mu \bigtimes_{t=1}^{N} J_{k_t}^t = \sum_{r=1}^{m} \left[\sum_{k_1=\alpha_{1r}}^{\beta_{1r}} \cdots \sum_{k_N=\alpha_{Nr}}^{\beta_{Nr}} \mu \bigtimes_{t=1}^{N} J_{k_t}^t \right] = \sum_{r=1}^{m} \mu I_r .$$

This completes the proof. ∎

As a corollary to the finite additivity property, the sufficiency of the condition in Theorem 2.10 will be established.

Proof: It is to be shown that if $\Delta_{1\ldots N}\, g(a;\, b - a) \geq 0$ for any $[a, b)$ in $\mathcal{B}_I$ and if $\lim_{a_t \to b_t} \Delta_{1\ldots N}\, g(a;\, b - a) = 0$ for $t = 1, 2, \ldots, N$ where $a_t \leq b_t$, then $\lim_{(a_{i_1}, \ldots, a_{i_k}) \to (b_{i_1}, \ldots, b_{i_k})} \Delta_{1\ldots N}\, g(a;\, b - a) = 0$ for any $\{i_1, \ldots, i_k\} \subset \{1, 2, \ldots, N\}$, $a_{i_k} \leq b_{i_k}$.

Due to Theorem 2.2 it is sufficient to consider the case where $(i_1, \ldots, i_k) = (1, 2, \ldots, k)$. Write $f(a_1, \ldots, a_k) = \Delta_{1\ldots N}\, g(a;\, b - a)$, where $a_{k+1}, \ldots, a_N$ and b are considered fixed. Consider a monotonic sequence $\left\{(a_1^n,\, a_2^n, \ldots, a_k^n)\right\}_{n=1}^{\infty}$ for which $a_i^n \leq b_i$ and $\lim_{n\to\infty} a_i^n = b_i$, $i = 1, 2, \ldots, k$. It will be shown that for any such sequence $\lim_{n\to\infty} f(a_1^n, a_2^n, \ldots, a_k^n) = 0$. This is sufficient to establish the theorem.

Consider the intervals

$$I_n = I_n^1 \times \bigtimes_{t=2}^{N} I^t$$

for $n = 1, 2, \ldots,$ where $I_n^1 = [a_1^n,\, b_1)$ and $I^t = [a_t^1, b_t)$. Also consider

$$J_n = I_n^1 \times \left(\bigtimes_{s=2}^{k} I_n^s \right) \times \left(\bigtimes_{t=k+1}^{N} I^t \right)$$

where $I_n^s = [a_s^n, b_s)$, $s = 1,2,\ldots,k$. Since $a_s^1 \leq \ldots \leq a_s^n \leq b_s$,

$$I^s = \sum_{k_s=1}^{n} J_{k_s}^s ,$$

where, for $k_s = 1,2,\ldots,n-1$, $J_{k_s}^s = [a_s^{k_s}, a_s^{k_s+1})$, and, for $k_s = n$, $J_{k_s}^s = [a_s^n, b_s) = I_n^s$. Thus

$$I_n = I_n^1 \times \left(\bigtimes_{s=2}^{k} \sum_{k_s=1}^{n} J_{k_s}^s \right) \times \left(\bigtimes_{t=k+1}^{N} I^t \right).$$

By Theorem 4.7

$$I_n = \sum \left[I_n^1 \times \left(\bigtimes_{s=2}^{k} J_{k_s}^s \right) \times \left(\bigtimes_{t=k+1}^{N} I^t \right) \right],$$

the summation being over $k_s = 1,2,\ldots,n$, for $s = 2,3,\ldots, k$.

The term corresponding to $k_2 = k_3 = \ldots = k_k = n$ is

$$I_n^1 \times \left(\bigtimes_{s=2}^{k} I_n^s \right) \times \left(\bigtimes_{t=k+1}^{N} I^t \right) = J_n .$$

Hence I_n is the disjoint union of J_n and $n^{k-1} - 1$ intervals in $\mathcal{B}_I$. It follows from the finite additivity of μ that

$$\mu J_n \leq \mu I_n .$$

Now let $\varepsilon > 0$ be given. By hypothesis, and since $\lim_{n\to\infty} a_1^n = b_1$, $a_1^n \leq b_1$, there is an n_ε such that

$$\Delta_{1\ldots N}\, g(a_1^n, a_2^1, \ldots, a_N^1 ; b_1 - a_1^n, b_2 - a_2^1, \ldots, b_N - a_N^1) < \varepsilon$$

whenever $n > n_\varepsilon$. This difference is μI_n. But

$$f(a_1^n, \ldots, a_k^n) = \mu J_n \leq \mu I_n .$$

This completes the proof of Theorem 2.10. ∎

It remains to prove that μ is σ-additive.

Theorem 4.9: The set function $\mu = \Delta_{1\ldots N} F$ is σ-additive on $\mathcal{B}_I$.

Proof: Let $I = [a, b) = \bigtimes_{t=1}^{N} [a^t, b^t) = \sum_{k=1}^{\infty} I_k$ where the $[a^t, b^t)$ are one-dimensional intervals and I and all I_k are in $\mathcal{B}_I$; let $I_{k_1}, I_{k_2}, \ldots, I_{k_m}$ be a finite subset of the I_k; and let $I_k = [a_k, b_k) = \bigtimes_{t=1}^{N} [a_k^t, b_k^t) = \bigtimes_{t=1}^{N} I_k^t$.

Again, for fixed t, use the ordering device on the $a^t, b^t, a_{k_r}^t, b_{k_r}^t$, $r = 1,2,\ldots,m$, to obtain $a^t = x_0^t \leq x_1^t \leq \ldots \leq x_{n_t}^t = b^t$. These orderings effect the usual grid-like decomposition of I. Furthermore, some of these intervals add to $\sum_{r=1}^{m} I_{k_r}$. By finite additivity, then,

$$\mu I \geq \sum_{r=1}^{m} \mu I_{k_r} ;$$

and since this is true for every subset $\{k_1, k_2, \ldots, k_m\}$,

$$\mu I \geq \sum_{k=1}^{\infty} \mu I_k .$$

It will now be shown that the reverse inequality also holds. Assume that I and all the I_k are not empty - for if I is empty there is nothing to prove, and, if some I_k is empty, it may be deleted. Define $\bar{I}^\delta = \bigtimes_{t=1}^{N} \bar{I}_\delta^t = \bigtimes_{t=1}^{N} [a^t, b^t - \delta]$ where $a^t \leq b^t - \delta$ for each t. Define $I^\delta = \bigtimes_{t=1}^{N} I_\delta^t = \bigtimes_{t=1}^{N} [a^t, b^t - \delta)$. For each such $\delta > 0$

$$I = \bigtimes_{t=1}^{N} (A_1^t + A_2^t) ,$$

where $A_1^t = [a^t, b^t - \delta)$ and $A_2^t = [b^t - \delta, b^t)$. By the distributive law

$$I = \bigtimes_{t=1}^{N} A_1^t + \sum \bigtimes_{t=1}^{N} A_{k_t}^t ,$$

where the summation is over $k_t = 1, 2$, $t = 1, 2, \ldots, N$, except $k_1 = k_2 = \ldots = k_N = 1$. Notice that $\bigtimes_{t=1}^{N} A_1^t = I^\delta$ and for every other term $\bigtimes_{t=1}^{N} A_{k_t}^t$ in the sum $k_t = 2$ for some t. So, by finite additivity,

$$\mu I = \mu I^\delta + \sum_{j=1}^{2^N - 1} \mu D_j^\delta ,$$

where the D_j^δ are the sets $\bigtimes_{t=1}^{N} A_{k_t}^t$ with at least one $k_t = 2$. But for each $j = 1, 2, \ldots, 2^N - 1$,

$$\mu D_j^\delta = \Delta_{1\ldots N} F(x_1, \ldots, x_N; y_1 - x_1, \ldots, y_N - x_N) ,$$

where $[x_t, y_t)$ is either $[a^t, b^t - \delta)$ or $[b^t - \delta, b^t)$. Also $x_t = b^t - \delta$ for some t. It follows that $y_t - x_t = \delta$ for some t. Thus by Theorem 2.10

$$\lim_{\delta \to 0^+} \mu D_j^\delta = 0 , \quad j = 1, 2, \ldots, 2^N - 1 .$$

Now let $\varepsilon > 0$ be given. For each j there is a δ_j such that

$$\mu D_j^\delta < \frac{\varepsilon}{2(2^N - 1)} , \quad j = 1, 2, \ldots, 2^N - 1,$$

whenever $\delta < \delta_j$. Let δ_0 be the smallest of the $\delta_1, \delta_2, \ldots, \delta_{2^N - 1}$,

$b^1 - a^1, \ldots, b^N - a^N$. It follows that for any $\delta < \delta_o$

$$\mu I = \mu I^{\delta} + \sum_{j=1}^{2^N-1} \mu D_j^{\delta} < \mu I^{\delta} + \frac{\varepsilon}{2} .$$

Now for an arbitrary k, let $I_k^{\delta_k} = \bigtimes_{t=1}^{N} I_k^{t,\delta_k}$, where $I_k^{t,\delta_k} = (a_k^t - \delta_k, b_k^t)$, $\delta_k > 0$, and let $\bar{I}_k^{\delta_k} = \bigtimes_{t=1}^{N} \bar{I}_k^{t,\delta_k} = \bigtimes_{t=1}^{N} [a_k^t - \delta_k, b_k^t)$. Then $\bar{I}_k^{\delta_k} = \bigtimes_{t=1}^{N} (B_{1k}^t + B_{2k}^t)$ where $B_{1k}^t = [a_k^t - \delta_k, a_k^t)$ and $B_{2k}^t = [a_k^t, b_k^t)$. Employing the same method used previously, it follows that

$$\bar{I}_k^{\delta_k} = I_k + \sum_{j=1}^{2^N-1} E_{jk}^{\delta_k} ,$$

where each $E_{jk}^{\delta_k}$ is one of the sets $\bigtimes_{t=1}^{N} B_{k_t k}^t = \bigtimes_{t=1}^{N} [x_{k_t}^t, y_{k_t}^t)$, and where, for some t, $[x_{k_t}^t, y_{k_t}^t) = [a_k^t - \delta_k, a_k^t)$. From this, from the finite additivity of μ, and from the left-continuity of F, there exists a number δ_k^o such that when $\delta_k < \delta_k^o$

$$\mu E_{jk}^{\delta_k} < \frac{\varepsilon}{2 \cdot 2^k (2^N - 1)} .$$

This inequality implies that, for each $k = 1,2,\ldots,$

$$\mu \bar{I}_k^{\delta_k} = \mu I_k + \sum_{j=1}^{2^N-1} \mu E_{jk}^{\delta_k} < \mu I_k + \frac{\varepsilon}{2^{k+1}} ,$$

for any fixed $\delta_k < \delta_k^o$. Thus for each $k = 1, 2, \ldots, \bar{I}_k^{\delta_k} \supset I_k^{\delta_k} \supset I_k$, so that

$$\bigcup_{k=1}^{\infty} \bar{I}^{\delta_k} \supset \bigcup_{k=1}^{\infty} I_k^{\delta_k} \supset \sum_{k=1}^{\infty} I_k = I \supset \bar{I}^{\delta} \supset I^{\delta} .$$

It follows that $\{I_k^{\delta_k}\}$ is an open cover of the compact set $\bar{I}^{\delta}$ so that by the Heine-Borel theorem there exist sets $I_{k_1}^{\delta_{k_1}}, \ldots, I_{k_M}^{\delta_{k_M}}$ such that

$$\bigcup_{r=1}^{M} \bar{I}_{k_r}^{\delta_{k_r}} \supset \bar{I}^{\delta} \supset I^{\delta} .$$

It will now be shown that $\sum_{r=1}^{M} \mu\, \bar{I}_{k_r}^{\delta_{k_r}} \geq \mu I^{\delta}$. The ordering device is used again. For fixed t consider the intervals $[a_{k_r}^t - \delta_{k_r}, b_{k_r}^t)$. For some r', $a_{k_{r'}}^t - \delta_{k_{r'}} < a^t$, and for some r'', $b_{k_{r''}}^t > b^t - \delta$. Order these numbers, including a^t and $b^t - \delta$, and label them

$$a_{k_{r'}}^t - \delta_{k_{r'}} = c_o^t \leq c_1^t \leq \ldots \leq c_{m_t}^t = b_{k_{r''}}^t .$$

Notice that $c_o^t < a^t$ and $c_{m_t}^t > b^t - \delta$. Denote $[c_{k_t-1}^t, c_{k_t}^t)$ by $J_{k_t}^t$ and $[c_o^t, c_{m_t}^t)$ by J^t.

Since for each r there exist a α_{rt} and a β_{rt} such that

$$[a_{k_r}^t - \delta_{k_r}, b_{k_r}^t) = \sum_{n_t=\alpha_{rt}}^{\beta_{rt}} J_{n_t}^t ,$$

it is the case that

$$\overline{I}_{k_r}^{\delta k_r} = \mathop{\times}_{t=1}^{N} [a_{k_r}^t - \delta_{k_r}, b_{k_r}^t) = \sum_{n_1=\alpha_{r1}}^{\beta_{r1}} \cdots \sum_{n_N=\alpha_{rN}}^{\beta_{rN}} \mathop{\times}_{t=1}^{N} J_{n_t}^t .$$

It follows from the finite additivity of μ that

$$\sum_{r=1}^{M} \mu \overline{I}_{k_r}^{\delta k_r} = \sum_{r=1}^{M} \sum_{n_1=\alpha_{r1}}^{\beta_{r1}} \cdots \sum_{n_N=\alpha_{rN}}^{\beta_{rN}} \mu \mathop{\times}_{t=1}^{N} J_{n_t}^t .$$

Since the a^t and $b^t - \delta$ appear in the lists of the c_k^t, the expression on the right is a sum of values of μ on intervals some of which add to I^δ. Another appeal to the finite additivity of μ yields that

$$\sum_{r=1}^{M} \mu \overline{I}_{k_r}^{\delta k_r} \geq \mu I^\delta .$$

Thus far it has been established that for any $\varepsilon > 0$

$$\mu I < \mu I^\delta + \frac{\varepsilon}{2} ; \ \mu I^\delta \leq \sum_{r=1}^{M} \mu \overline{I}_{k_r}^{\delta k_r} \leq \sum_{k=1}^{\infty} \mu \overline{I}_k^{\delta k} ; \text{ and } \mu \overline{I}_k^{\delta k} < \mu I_k + \frac{\varepsilon}{2^{k+1}} .$$

Hence for any $\varepsilon > 0$

$$\mu I < \mu I^\delta + \frac{\varepsilon}{2} \leq \sum_{k=1}^{\infty} \mu \overline{I}_k^{\delta k} + \frac{\varepsilon}{2} < \sum_{k=1}^{\infty} (\mu I_k + \frac{\varepsilon}{2^{k+1}}) + \frac{\varepsilon}{2} = \sum_{k=1}^{\infty} \mu I_k + \varepsilon .$$

This establishes the reverse inequality and completes the proof of the theorem. ∎

Part II has now been established.

The Proof of Part III

It is sufficient, as in the one-dimensional case, to exhibit a distribution function F for which $\Delta_{1\ldots N} F = \mu$ for any given $\mu \in \mathcal{M}(\mathcal{B}_I)$.

The technique employed is to represent any interval in $\mathcal{B}_I$ in terms of intervals one of whose vertices is the origin. F is then defined in terms of the measure μ evaluated at such intervals. It is instructive to consider the construction of F in E^2.

If $[a_o, a_1)$ is an interval in E^2 where $a_j = (a_j^1, a_j^2)$, $j = 1, 2$, which lies in the first quadrant it can be written as follows:

$$[a_o, a_1) = [o, a_1^1) \times [o, a_1^2) - [o, a_o^1) \times [o\ a_1^2) +$$
$$- \{[o, a_1^1) \times [o, a_o^2) - [o,\ a_o^1) \times [o,\ a_o^2)\}.$$

If $[a_o, a_1)$ lies in the first and fourth quadrants it can be represented as

$$[a_o, a_1) = [o,\ a_1^1) \times [a_o^2, o) + [o,\ a_1^1) \times [o,\ a_1^2) - [o,\ a_o^1) \times [a_o^2, o) +$$
$$- [o,\ a_o^1) \times [o,\ a_1^2).$$

The other cases are treated similarly. Each case is subsumed under the following general formula:

$$[a_o,\ a_1) = \bigtimes_{t=1}^{2} \left\{(-1)^{s_o^t + 1} [m_o^t, M_o^t) + (-1)^{s_1^t} [m_1^t, M_1^t)\right\}$$

where $m_i^t = \min \{0,\ a_i^t\}$, $M_i^t = \max \{0, a_i^t\}$, and s_i^t is 0 in case $m_i^t = 0$ and 1 otherwise. That the notation is meaningful is readily verified.

It remains to define F on E^2. A little experimentation suggests that, if $F(0) = 0$, then

$$F(x) = (-1)^s \mu[(m_1, m_2), (M_1, M_2)),$$

where $m_t = \min\{0, x_t\}$ and $M_t = \max\{0, x_t\}$, will be a distribution function. Here the exponent is the number of nonzero m_t. It is easily seen that the interval on which μ is defined has a vertex at the origin. That F is a distribution function follows from a subsequent theorem for the general case in E^N.

The procedure for the general case is analogous except that the notions of indicator functions and expected value are explicitly used. It is contained in what follows.

<u>Lemma 4.10</u>: Let $a_o = (a_o^1, \ldots, a_o^N)$ and $a_1 = (a_1^1, \ldots, a_1^N)$. Then

$$\Delta_{1\ldots N} F(a_o; a_1 - a_o) = \sum_{k=o}^{N} (-1)^k \sum_{\sum_1^N i_t = N-k} F(a_{i_1}^1, \ldots, a_{i_N}^N).$$

<u>Proof</u>: This is a special case of Theorem 2.4.

<u>Lemma 4.11</u>. Let $[a_o, a_1) = \bigtimes_{t=1}^{N} [a_o^t, a_1^t)$. Let $m_i^t = \min\{o, a_i^t\}$, $M_i^t = \max\{o, a_i^t\}$, $A_o^t = [m_o^t, M_o^t) \times \left(\bigtimes_{t' \neq t} E_{t'}\right)$, and $A_1^t = [m_1^t, M_1^t) \times \left(\bigtimes_{t' \neq t} E_{t'}\right)$. Define s_i^t to be 0 if $m_i^t = 0$ and 1 otherwise so that $\sum_{i=1}^{N} s_i^t$ is the number of nonzero m_i^t for given t. Then

$$[a_o^t, a_1^t) = (-1)^{s_o^t + 1} [m_o^t, M_o^t) + (-1)^{s_1^t} [m_1^t, M_1^t),$$

and

$$[a_0, a_1) = \bigcap_{t=1}^{N} [(-1)^{s_0^t+1} A_0^t + (-1)^{s_1^t} A_1^t].$$

Proof: It is easily seen that if $[a_0^t, a_1^t)$ is not empty it is impossible for s_0^t to be 0 and, at the same time, for s_1^t to be 1, and that either $[m_0^t, M_0^t)$ and $[m_1^t, M_1^t)$ are disjoint, or one is a subset of the other. The cylinders A_0^t and A_1^t behave the same way. The notation therefore is appropriate.

The set-theoretic identities follow from familiar arguments. ∎

It is helpful to use indicator functions and the notion of expected value.

Definition 4.12. Let $A \subset E^N$. The function I_A which is zero on A and unity otherwise is called the indicator function of A.

Lemma 4.13. $I_{\bigcap_{j=1}^{m} A_j} = \prod_{j=1}^{m} I_{A_j}$, $I_{\sum_{j=1}^{m} A_j} = \sum_{j=1}^{m} I_{A_j}$, and $I_{A_1-A_2} = I_{A_1} - I_{A_2}$ if $A_2 \subset A_1$.

Proof: The proof is obvious. ∎

Definition 4.14: Suppose $E^N = \sum_{j=1}^{m} A_j$ and that $f(x) = \sum_{j=1}^{m} f_j I_{A_j}(x)$ where each f_j is a number. Then the expected value $E[f]$ of f relative to a given measure μ is defined to be $\sum_{j=1}^{m} f_j \mu A_j = \int f \, d\mu$.

Lemma 4.15: $E\left[\sum_{j=1}^{m} c_j I_{A_j}\right] = \sum_{j=1}^{m} c_j \mu A_j$.

Proof: The proof is straightforward and is omitted. ∎

Theorem 4.16:

$$I_{[a_o,a_1)} = \sum_{k=o}^{N} (-1)^k \sum_{\Sigma i_t = N-k} (-1)^{\Sigma s_{i_t}^t} I_{[m_{i_1\ldots i_N}, M_{i_1\ldots i_N})}$$

where $m_{i_1\ldots i_N} = (m_{i_1}^1, m_{i_2}^2, \ldots m_{i_N}^N), M_{i_1\ldots i_N} = (M_{i_1}^1, M_{i_2}^2, \ldots, M_{i_N}^N)$ and the $s_{i_t}^t$, $m_{i_t}^t$, and $M_{i_t}^t$ are as in Lemma 4.11.

Proof: By the second assertion of Lemma 4.11

$$I_{[a_o, a_1)} = I_{\bigcap_{t=1}^{N} [(-1)^{s_o^t+1} A_o^t + (-1)^{s_1^t} A_1^t]} .$$

By Lemma 4.13, the right side can be written as $\prod_{t=1}^{N} [(-1)^{s_o^t+1} I_{A_o^t} + (-1)^{s_1^t} I_{A_1^t}]$

The distributive law for real numbers allows this to be written

$$\sum_{k=o}^{N} \sum_{\Sigma i_t = N-k} (-1)^k (-1)^{\Sigma s_{i_t}^t} \prod_{t=1}^{N} I_{A_{i_t}^t} .$$

Now apply Lemma 4.13 to $\prod_{t=1}^{N} I_{A_{i_t}^t}$ and notice that $\bigcap_{t=1}^{N} A_{i_t}^t =$ $= \times_{t=1}^{N} [m_{i_t}^t, M_{i_t}^t)$. The conclusion follows. ∎

The next theorem is the culmination of this section.

Theorem 4.17: Let μ be a given Lebesgue-Stieltjes measure on $\mathcal{B}_I$.

Consider $F(x) = (-1)^s \mu[m, M)$ for $x \neq 0$, $F(0) = 0$, where the components m_t and M_t of m and M are $m_t = \min\{0, x_t\}$ and $M_t = \max\{0, x_t\}$ with $x = (x_1, x_2, \ldots, x_N)$, and where s is the number of nonzero m_t. Then $\Delta_{1\ldots N} F(a_o; a_1 - a_o) = \mu[a_o, a_1)$ for all $[a_o, a_1) \varepsilon \mathcal{B}_I$, and F is a distribution function on E^N.

<u>Proof</u>: First apply the definition of F to the right side of the equation in Lemma 4.10. Then apply Lemma 4.15 to the result to obtain

$$\Delta_{1\ldots N} F(a_o; a_i - a_o) = \sum_{k=o}^{N} (-1)^k \sum_{\Sigma i_t = N-k} F(a^1_{i_1}, a^2_{i_2}, \ldots, a^N_{i_N}) =$$

$$= \sum_{k=o}^{N} (-1)^k \sum (-1)^{\Sigma s^t_{i_t}} \mu[m_{i_1\ldots i_N}, M_{i_1\ldots i_N}) =$$

$$= \sum_{k=o}^{N} (-1)^k \sum (-1)^{\Sigma s^t_{i_t}} E\left[I_{[m_{i_1\ldots i_N}, M_{i_1\ldots i_N})}\right] =$$

$$= E\left[\sum (-1)^k \sum (-1)^{\Sigma s^t_{i_t}} I_{[m_{i_1\ldots i_N}, M_{i_1\ldots i_N})}\right].$$

By Theorem 4.16 the last expression is $E\left[I_{[a_o, a_1)}\right] = \mu[a_o, a_1)$.

It remains to prove that F is a distribution function. Clearly F is finite and $\Delta_{1\ldots N} F \geq 0$ since μ is a measure. Since

$$\lim_{a_o^t \to a_1^t -} \Delta_{1\ldots N} F(a_o; a_1 - a_o) = \lim_{a_o^t \to a_1^t -} \mu[a_o, a_1) = \mu \lim_{a_o^t \to a_1^t -} [a_o, a_1) = 0$$

for $t = 1, 2, \ldots, N$, F is continuous from the left by Theorem 2.10. This completes the proof. ∎

Theorem 4.17 gives the distribution function sufficient for the proof of Part III. The proof of the correspondence theorem in E^N is now complete.

CHAPTER V

SOME GENERALIZATIONS

The class $\mathcal{B}_I$ in E^N is not a field and it is therefore difficult to use the version of the Carathéodory extension theorem stated in the introduction to prove Part I. However, the class $\mathcal{B}_I$ is a half-ring; and Part I of the correspondence theorem can be deduced from results in VON NEUMANN and HALMOS. Theorems 10.1.12 and 10.1.2 in VON NEUMANN state that a measure μ on a half-ring $\mathcal{R}$ extends uniquely to the minimal ring $\mathcal{R}_m$ over $\mathcal{R}$ and that $\mathcal{R}_m$ is the collection of all finite disjoint unions of sets in $\mathcal{R}$. A stronger form of the extension theorem is found in Theorem A, § 13, of HALMOS. It states that μ on $\mathcal{R}_m$ extends uniquely to the minimal σ-ring $\mathcal{B}_m$ over $\mathcal{R}_m$. In the case of the correspondence theorem $\mathcal{B}_m$ is the Borel field $\mathcal{B}$. These results prove Part I up to the extension to the completion $\mathcal{B}_\mu$ of the Borel field for μ.

A portion of Part II and all of Part III can be obtained from other results in VON NEUMANN, pp. 160-167. These results are stated in Theorems 5.3 - 5.6 with slight modifications and without proof. Theorem 5.7 is also from VON NEUMANN and relates intimately to the correspondence theorem.

<u>Definition 5.1</u>: Denote $[a, b)$ where $a = (a_1,\ldots,a_N)$ and $b = (b_1,\ldots,b_N)$ by $I\begin{bmatrix} a_1,\ldots,a_N \\ b_1,\ldots,b_N \end{bmatrix}$. Since I is either empty - in which case $a_t = b_t$ for at least one t - or all a_t and b_t satisfy $a_t < b_t$ and are uniquely determined by I, any set function μ on $\mathcal{B}_I$

which vanishes at the empty set can be thought of as a point function $F_\mu \begin{bmatrix} a_1,\ldots,a_N \\ b_1,\ldots,b_N \end{bmatrix}$ of the variables $a_1,\ldots,a_N$, $b_1,\ldots,b_N$ which is defined for $a_t \leq b_t$ and which has the value zero if some $a_t = b_t$. That is,

$$\mu I \begin{bmatrix} a_1,\ldots,a_N \\ b_1,\ldots,b_N \end{bmatrix} = F_\mu \begin{bmatrix} a_1,\ldots,a_N \\ b_1,\ldots,b_N \end{bmatrix}.$$

F_μ is said to be the <u>point function belonging to μ</u>.

<u>Definition 5.2</u>: A real valued function μ on a half-ring $\mathcal{R}$ is called a <u>finite measure function</u> in case μ is non-negative, finite, and finitely additive on $\mathcal{R}$. It is assumed that the original space is covered by a sequence of sets of $\mathcal{R}$.

<u>Theorem 5.3</u>: Suppose μ is a finite measure function on the half-ring $\mathcal{B}_I$. Then F_μ satisfies

$$F_\mu \begin{bmatrix} a_1,\ldots,a_N \\ b_1,\ldots,b_N \end{bmatrix} = F_\mu \begin{bmatrix} a_1,\ldots,a_t,\ldots,a_N \\ b_1,\ldots,e_t,\ldots,b_N \end{bmatrix} + F_\mu \begin{bmatrix} a_1,\ldots,e_t,\ldots,a_N \\ b_1,\ldots,b_t,\ldots,b_N \end{bmatrix}$$

where $a_\sigma \leq b_\sigma$, $\sigma = 1,2,\ldots,N$, and $a_t \leq e_t \leq b_t$. Conversely any finite F satisfying this equation determines a finite measure function μ. Moreover, the function

$$F_\mu^o \begin{bmatrix} a_1,\ldots,a_N \\ b_1,\ldots,b_N \end{bmatrix} = (-1)^e \, F_\mu \begin{bmatrix} \hat{a}_1,\ldots,\hat{a}_N \\ \hat{b}_1,\ldots,\hat{b}_N \end{bmatrix}$$

where $\hat{a}_\sigma = \min \{a_\sigma, b_\sigma\}$, $\hat{b}_\sigma = \max \{a_\sigma, b_\sigma\}$, and e is the number of indices for which $a_\sigma > b_\sigma$, is the unique extension of F_μ to all of E^{2N} for which the equation holds for all e_t.

Theorem 5.4: F_μ^o is the extended function of some finite measure function on $\mathcal{B}_I$ in the sense of the previous theorem if and only if there is a finite real-valued function φ on E^N for which

$$F_\mu^o \begin{bmatrix} a_1,\dots,a_N \\ b_1,\dots,b_N \end{bmatrix} = \sum_{r_1,\dots,r_N=1,2} (-1)^{r_1+\dots+r_N} \varphi(c_1^{(r_1)},\dots,c_N^{(r_N)})$$

where $c_\sigma^{(1)} = a_\sigma$ and $c_\sigma^{(2)} = b_\sigma$, $\sigma = 1,2,\dots,N$.

Theorem 5.5. The extended function $F_\mu^o \begin{bmatrix} a_1,\dots,a_N \\ b_1,\dots,b_N \end{bmatrix}$ does not determine φ. But if φ is one such function, φ' is another if and only if there are N real-valued functions $\Psi_1,\dots,\Psi_N$ defined on E^{N-1} such that

$$\varphi'(c_1,\dots,c_N) = \varphi(c_1,\dots,c_N) + \sum_{t=1}^{N} \Psi_t(c_1,\dots,c_{t-1},c_{t+1},\dots,c_N).$$

Moreover there is a unique φ such that $\varphi(c_1,\dots,c_N) = 0$ if any $c_\sigma = 0$. This unique φ is

$$\varphi(c_1,\dots,c_N) = F_\mu^o \begin{bmatrix} 0,\dots,0 \\ c_1,\dots,c_N \end{bmatrix}.$$

Theorem 5.6: A finite measure function μ on $\mathcal{B}_I$ is non-negative and totally additive if and only if at least one of its φ's has the monotone property of Theorem 2.10 and is continuous from the left.

Theorem 5.7: Let $\mathcal{R}$ be a half ring in a topological space S. A finite measure function μ on $\mathcal{R}$ is σ-additive on $\mathcal{R}$ if for every $\varepsilon > 0$ and every $M \in \mathcal{R}$, there exist sets P and Q in $\mathcal{R}$, a compact set C

and an open set G such that $P \subset C \subset M \subset G \subset Q$ and such that $\mu P \geq \mu M - \varepsilon$ and $\mu Q \leq \mu M + \varepsilon$.

Proof: It is sufficient to establish that if $M = \sum_{i=1}^{\infty} N_i$ where M, $N_1, N_2, \ldots$ are all in $\mathcal{R}$, then $\mu M \leq \sum_{i=1}^{\infty} \mu N_i$ since, for any n, $\sum_{i=1}^{n} \mu N_i = \mu \sum_{i=1}^{n} N_i \leq \mu M$ so that the reverse inequality $\sum_{i=1}^{\infty} \mu N_i \leq \mu M$ holds.

Let $\varepsilon > 0$ be given. There exist a set $P \in \mathcal{R}$ and a compact set C such that

$$\mu P \geq \mu M - \frac{\varepsilon}{2} .$$

Since each $N_i \in \mathcal{R}$, there are sets $Q_i \in \mathcal{R}$ and open sets G_i such that $N_i \subset G_i \subset Q_i$ and

$$\mu Q_i \leq \mu N_i + \frac{\varepsilon}{2^{i+1}} .$$

Since $C \subset M \subset \sum_{i=1}^{\infty} N_i \subset \sum_{i=1}^{\infty} G_i$, the sets G_i form an open cover of C. By the Heine-Borel theorem there is an n such that

$$C \subset \sum_{i=1}^{n} G_i .$$

Thus $P \subset C \subset \sum_{i=1}^{n} G_i \subset \sum_{i=1}^{n} Q_i$ so that

$$\mu M \le \mu P + \frac{\varepsilon}{2} \le \sum_{i=1}^{n} \mu Q_i + \frac{\varepsilon}{2} \le \sum_{i=1}^{\infty} \mu Q_i + \frac{\varepsilon}{2} \le \sum_{i=1}^{\infty} \mu N_i + \frac{\varepsilon}{2^{i+1}} + \frac{\varepsilon}{2} \le \sum_{i=1}^{\infty} \mu N_i + \varepsilon .$$

Since ε is arbitrary, $\mu M \le \sum_{i=1}^{\infty} \mu N_i$. This completes the proof. ∎

Theorem 4.9 is proved as a special case of the previous theorem with I, I_i, I^δ, $\bar{I}^\delta$, $\bar{I}_k^{\delta_k}$, and $I_k^{\delta_k}$ playing rôles similar to those of M, N, P, C, Q_i and G_i, respectively.

The sufficiency of the condition on φ in Theorem 5.6 rests on constructing sets P, C, G, and Q meeting the conditions in Theorem 5.7.

Since φ is finite the conditions on φ in Theorem 5.6 say that φ is a distribution function. Thus a paraphrase of Theorem 5.6 is that a finite measure function μ on $\mathcal{B}_I$ is also a measure on $\mathcal{B}_I$ if and only if one of its φ's is a distribution function.

Theorem 5.4 states that

$$\mu[a,b) = \sum_{r_1,\ldots,r_N=1,2} (-1)^{r_1+\ldots+r_N} \varphi\left(c_1^{(r_1)},\ldots,c_N^{(r_N)}\right).$$

The proof that this sum is $\Delta_{1\ldots N}\,\varphi(a;\, b - a)$ parallels that of Theorem 2.4.

Since all the Ψ_t are defined on E^{N-1}, the condition on φ in Theorem 5.5 implies that $\Delta_{1\ldots N}\,\varphi = \Delta_{1\ldots N}\,\varphi'$. Conversely, von Neumann shows in his proof of Theorem 5.5 that if $\Delta_{1\ldots N}\,\varphi' = \Delta_{1\ldots N}\,\varphi$, then

$\varphi' = \varphi + \sum_{i=1}^{N} \Psi_t$. It follows that the class of all φ which generate the same extended function F_{μ}^{o} is the equivalence class described in Part II containing one such φ. Thus, as soon as it is known that $\mu = \Delta_{1\ldots N}F$ is finitely additive on $\mathcal{B}_I$, Theorem 5.6 implies that μ is σ-additive.

It has been stated that to prove Part III it is sufficient to demonstrate the existence of a distribution function F for which $\mu = \Delta_{1\ldots N}F$. While there is still interest in such a direct construction Theorem 5.6 guarantees the existence of one and thus establishes Part III without the actual construction. Furthermore, it is easily seen that the distribution function constructed in Chapter IV is the unique one mentioned in Theorem 5.5.

BIBLIOGRAPHY

CRAMÉR, H., [1] Mathematical Methods of Statistics, Princeton: Princeton University Press, 1946.

CRAMÉR, H., [2] Random Variables and Probability Distributions, Cambridge: Cambridge Tracts In Mathematics and Mathematical Physics, No. 36, 1937.

HALMOS, P. R., Measure Theory, New York: D. Van Nostrand Co., 1950.

KAMKE, E., Das Lebesgue-Stieltjes Integral, Leipzig: B. G. Teubner Verlagsgesellschaft, 1956.

KOLMOGOROV, A. N., Foundations of the Theory of Probability, New York: Chelsea Publishing Co., 1950.

LEBESGUE, H., Leçons sur Intégration et la Recherche des Fonctions Primitives, 2-me ed., Paris: Gauthier-Villars et Cie., 1928.

LOÈVE, M., Probability Theory, Princeton, N. J.: D. Van Nostrand Co. Inc., 1960.

VON NEUMANN, J., Functional Operators I, Princeton: Annals of Mathematics Studies, Vol. 21, 1950.

RADON, J., "Theorie und Andwendung der absolut additiven Mengenfunktionen," Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Klasse der kaiserlichen Akademie der Wissenschaften. Wien., Vol. 122, 1913, pp. 1295-1438.

DE LA VALLÉE POUSSIN, C., Intégrales de Lebesgue, fonctions d'Ensemble, Classes de Baire, Paris: Gauthier-Villars, 1934.

WILKS, S. S., [1] Mathematical Statistics, Princeton: Princeton University Press, 1943.

WILKS, S. S., [2] Mathematical Statistics, New York: John Wiley and Sons, Inc., 1962.